

Result No.	Query Match	Score	Length	DB	ID	Description
---------------	----------------	-------	--------	----	----	-------------

Direct Submission
Submitted (13-MAY-1998) Osamu Katoch, Res. Inst. Radiation Biology
and Medicine, Hiroshima Univ., Department of Environment and
Mutation, Kasumi 1-2-3, Minami-Ku, Hiroshima, Hiroshima 734-8553,

QY 1501 cacactcaagggttaaaccttatctctctctgagtcgagcgagctgtcttaagcctgaagt 1560
 |||||||
 Db 1501 CACACTCAGGGGTTAAACCTTATGTCTGCTGGAGTGGGGCACTGCTTTAGCCTGAAGT 1560
 QY 1561 caaaccttaacaacacacacagagtcacacacgagggagagccattgtatgtlaagagt 1620
 |||||||
 Db 1561 CAACCTTAAACAACACACAGAGGTGACACACGGGGGAGAGCCATTGTATGTAGCGAGT 1620
 QY 1621 gtggcgagagcttaccgggaatcaaccctgtatcaacgacacagaggaacacacacaggg 1680
 |||||||
 Db 1621 GTGGGGCAGAGCTTAAACCGGAAATCAACCTGATCAGCAGCAGACACACACTCAGGGG 1680
 QY 1681 agaaacattgtatgtctgagtcgagtcgaggaagagccttaagttaagtcacacccctt 1740
 |||||||
 Db 1681 AGAAGCATTGTATGTCTGCTGAGTGTGAGCAGAGCCTTTAATGATTAATCCACTCATTT 1740
 QY 1741 cacaccagagac 1800
 |||||||
 Db 1741 CACACACAGAGACACATTTCAGGGGAAAGCCTTTATGTGAGGAGGTGAGGAGAGGT 1800
 QY 1801 ttgcgcgaagccttaacctgtttaggcacacagagggcacacacacacacacacacacac 1860
 |||||||
 Db 1801 TTGCGCAGAAAGCCTTAACCTGTTTAGGCACAAAGGGCAGACACTCAGGTGCTTGTGTGCA 1860
 QY 1861 gggagctgggcaagccttctgtctgaagtaactctcatcaacacacacacacacacacac 1920
 |||||||
 Db 1861 GGGAGCTGGGCAAGGCTTTTGTCTGAAGTTAATCTCATTAACACACAGAGACACACAG 1920
 QY 1921 caagggggaagcctcatctgtctgcagagagtcgaggaagccttaagcgcagtcacacac 1980
 |||||||
 Db 1921 CAGGGGGAAAGCCTCATGTGTGCGAGGAGTGTGGCAAGGCTTATAGCCGCGACGTACACAC 1980
 QY 1981 tcattaac 2040
 |||||||
 Db 1981 TCATTAAACACACAGAGACACATTTCAGAGAGAGAAAGCCTTAATTTTGCAGAAAGTGTGAC 2040
 QY 2041 ggggctttagtcggaagtcacaacacacacacacacacacacacacacacacacacacac 2100
 |||||||
 Db 2041 GGGGCTTTAGTCGGAAGTCCACACTTATACAGACATCAGAGACACACACTCAGATAGAAG 2100
 QY 2101 ttatgtgtataggaagtgtgtacagcctttagcaggaagtcacttaactcaacacacac 2160
 |||||||
 Db 2101 TTATGTGTATAGGAATGTGTGTCACGCTTTTACCGAGATCATCTTCATCTCAGACACCC 2160
 QY 2161 agaaagac 2220
 |||||||
 Db 2161 AGAGGACACACACAGTGTGTGTGCTTTTTCAGCCATTGCTAGATACCAAGTGGAGACAT 2220
 QY 2221 tctgtgtatgattatgcatgagactgtactgtactgtactgtactgtactgtactgtactgt 2280
 |||||||
 Db 2221 TCTGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
 QY 2281 gagaatgtgtgtcatttcaagagacccctgtccctcccaactgtgagtggtgtgtgtgt 2340
 |||||||
 Db 2281 GAGATATGTGTGCTGCTATTTTCAGAGAGCCCTGCTCTCTCACTGTGATGTGGGTGTG 2340
 QY 2341 gaaacccggttcaagtaagtgaatgagagagagagagagagagagagagagagagagagag 2400
 |||||||
 Db 2341 GAAACCCGGGTCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
 QY 2401 gggactctgtgaaac 2460
 |||||||
 Db 2401 GGGTACTGTGTGAAACCAACCTTAAGCTGAAGACAGAGTCCGGGTAAATCTCTACTACTG 2460
 QY 2461 aatgagaacctgtctcccaacttgggtgtcttccctcccaactgtatcccaacccctcaac 2520
 |||||||
 Db 2461 AATGAGAACCTGTCTTCCCTTTGGTGTCTTCTCGATGATTCACCAACCTTTCACC 2520
 QY 2521 tatttcgtaactctcccttcccttaattgttttaccatgtctgtccacaccccttga 2580
 |||||||
 Db 2521 TATTTTACGTATACCTGCTCTTCTTCTTAATGTGTTTACACTGCTGTGCCACACCTTTTGA 2580

QY 2581 gtgtgtccttgcataacttaacaatcagtcacagctgtatctccctattctgagccataa 2640
 |||||||
 Db 2581 GTGTGCTCTTGTGCTACTTACAAATCAGTCAACTGTATTCCTTATTTCTAGGCCATTA 2640
 QY 2641 aagaccagactcagctgcagtcgagagagaaatcacccctgtctgtgaggttggggacca 2700
 |||||||
 Db 2641 AAGACCCAGACTCAGTGTGAGTGTGAGAGAGAAATCACCTGCTGTGAGGTTGGGGACCA 2700
 QY 2701 ctccctgcatccctctcccaactgagagctgtcttctgtcacaataaattcttctacc 2760
 |||||||
 Db 2701 CTCCCTGATCCCTCTCTCCACTGAGAGCTGTCTTTTGTCTCAATTAATTTCTTACCC 2760
 QY 2761 catccacacct 2772
 |||||||
 Db 2761 CATCCTCACCT 2772

RESULT 2
 AK056243 2854 bp mRNA linear PRI 31-OCT-2001
 LOCUS Homo sapiens cDNA FLJ131681 f1s, clone NT2R12005315, highly similar
 DEFINITION to Homo sapiens mRNA for HKR1.
 ACCESSION AK056243
 VERSION AK056243.1 GI:16551590
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens testocarcinoma cell line:NT2 cDNA to mRNA,
 clone 1lb:NT2R12 clone:NT2R12005315.

ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Ihibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
 Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
 Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
 Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
 Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawal,Hio,Y.,
 Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
 Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahari,K., Masuno,Y., Negai,K. and Isogai,T.

TITLE NEBO human cDNA sequencing project

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2854)
 AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
 TITLE Direct Submission

COMMENT Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Katsarazu, Chiba 297-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
 NEBO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.

FEATURES
 location/Qualifiers

1..2854
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="NT2R12005315"
 /cell_line="NT2"
 /cell_type="testocarcinoma"
 /note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
 precursor cells treated 2-weeks mitotic inhibitor after
 5-weeks retinoic acid (RA) induction.-majority NT2 neuron"
 BASE COUNT 761 a 692 c 754 g 647 t
 ORIGIN

Query Match 86.0%; Score 2383; DB 9; Length 2854;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 2433: Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
Qy	334	aaatccacatcttcaaacacaaactcattctcagctgagcgagggagcccttga	393	
Db	421	AAATTCATCTTCTTAACCAAAACATCATGTCTAGCTGAGGAGGGAAGGCCCTGGA	480	
Qy	394	gagagagagagaaatgctcacttgcaccttctccagaatcgagccgaatcaacta	453	
Db	481	GAGAGAGAGAAATGTCTCAGTGCACCTCTGTCCAGAAATCGAAGCCAGAAATTCACCTTA	540	
Qy	454	gtccctccctgcccctgacttcttccagtcagcagccttcagccaaatgctgctga	513	
Db	541	GTCCCTCCCTGCCCCCTGTGATTTTCTCCAGTCACAGACTCTCAGCCAACTGTGTGCTGA	600	
Qy	514	gtcactctctcagctgcttltcaaglttaaggcaggaatccctcctccctgggaaac	573	
Db	601	GTCATCTCTCTCAGCTGTCTTCAAGTTATGGGCAAGAAATCCTCTCCACCTGGGGAAC	660	
Qy	574	actatccagaagatcaagaacaacagcagatccattctgctttagtggcaagcagaat	633	
Db	661	ACTATCCAGAAATCAGAAACAAACAGCAGATCCATTTGCTTGTGTGCAAAAGCAGAAAT	720	
Qy	634	ggattcaagaaggagaagatccagatccctctgtttggagagtaagcaaaaatgacatt	693	
Db	721	GGATTCAAGAGGGAAGAACTCCAGACTCTCTTTGGAGATTAAGCAAAATGGCATTT	780	
Qy	694	caaaagcacttccagccacccctgaagaacaacagccagcacagtcaccaagagaaca	753	
Db	781	CAAAAGCACTTTCAGCCCACTGAAGAACACAGCCACAGTCACAGTCCAAAGGAACAAACA	840	
Qy	754	cagtggtgataataggtccagccctgaaacggagggcagatctagaaggaacaagcaag	813	
Db	841	CAGTGGTGGATATAGGGTCCAGCCCTGAAACGAGGGCAGCATCTTAAGGAACAGCAAAAG	900	
Qy	814	tattgcagtgttaagaatctcaggatttggagaatcaaatatagaaggtttggccag	873	
Db	901	TATTGCATGTTTAAAGTCTTCAGATTTTGGAAATCAAAATTAAGATTTGGGCCAG	960	
Qy	874	gcttatacgaagtcacaaactcctcttaagcctccagaagaacacaaactgggagacactt	933	
Db	961	GCTTATCAAGAGAGTCAAAACCTCTTAGCTCCAGAACACAAACTGGGAGACACTTT	1020	
Qy	934	acatgtacacagatgggggagagacacttggagatagtcagttcctcacaaccca	993	
Db	1021	ACATGTACACTGAGTGGGAGACACTTTGGCATGTGCACTCTCATCAAAAACCCAA	1080	
Qy	994	ggacacactctgggggaaagccttactgtgcaagggaatgtggcgagcctttagctga	1053	
Db	1081	GGACACACTCTGGGGGAAAGCCTTATGTGTGCAAGGAATGTGGCGAGCTTTAGCTGA	1140	
Qy	1054	agtcacacctgatacacatcagagagacacactcaggggagaacactatgtgtgcaag	1113	
Db	1141	AGTCAAACCTGTATCAACATCAGAGACACACTCAGGGGAGAAACCTTATGTGTCAAGG	1200	
Qy	1114	atttggagcagggcttctcttgggaagtcgaacctcttacaataagcggagaacctaag	1173	
Db	1201	ATTGTGGAGAGGCTTTACTTGTGAAGTGAACCTCTTTACATACATAGCGGAGACACTAG	1260	
Qy	1174	ggctcaagccttatgtgtgcaaggaatgtggcagagcctttagcctgaagtcaaacctca	1233	
Db	1261	GGCTCAAGCCTTATGTGTGCAAGGAATGTGGGCAAGCTTTAGCTGAAGTCAAAACCTCA	1320	
Qy	1234	ttaccaccagaaggcgacacacttgggagaagccttactgttgcagggaaatgtggcgctg	1293	
Db	1321	TTACCCACAGGAGGCGCACACTGGGAGAAAGCCTTATGTGTGAGGGAATGTGGCGG	1380	
Qy	1294	gcttcgcagcattcacacacctgtgcagacaaagaagacacatactcagagaggaagcctt	1353	
Db	1381	GCTTTCGCAGCATTCACACCTGTGCAAGACACAAAGGACACATTCAGAGAGAGCCCTT	1440	
Qy	1354	acatttgcagggagtgtagcaagcctttagcagaagatcacactctcaaacactcaa	1413	
Db	1441	ACATTTTGAGGAGAGTGTGAGCAAGGCTTTAGCCACAAAGTCAACCTCATCAGACACTTA	1500	
<hr/>				
Qy	1414	ggacacacacagagagaagaccttatgtatgcacagaatgtggcgctcaactttagctga	1473	
Db	1501	GGACACACACAGAGAGAAAGCCTTATGTATGACAGAAATGTGGGCTCACTTATAGCTGA	1560	
Qy	1474	aatcaaacctcaaaaacacacacagagaacacactcaggggttaaaccttatgtctgcctg	1533	
Db	1561	AATCAAAACCTCAAAACACACACAGAGACACACTCAGGGGTTAAACCTTATGTCTGCTGG	1620	
Qy	1534	agtcgagagatgctttagccttgaagtcgaacttaacaaacacagaggtacacacag	1593	
Db	1621	AGTGGGAGATGCTTTAGCTTGAAGTCAAACTTTAACAACACAGAGGTACACACGG	1680	
Qy	1594	gggagaagccattgtlatgcagagtgtyggcgaggtcttaaccggaatcaaaccttga	1653	
Db	1681	GGGAGAGCCATTGTATGTATGAGAGTGTGGCGAGGCTTTACCCGAAATCAACCTGA	1740	
Qy	1654	tcaagcaccagaagacacacactcaggggagaagcattgtatgtctgagtgtagcag	1713	
Db	1741	TCACGCACCCAGAGGACACACTCAGGGGAGAACCTTTGTATGTCTGTGAGTGAGCAG	1800	
Qy	1714	gcttataatgaatgccaccctcatttcaaccacagaagacacattcaggggaaagcctt	1773	
Db	1801	GCTTAAATGATTAAGTCCACCTCATTTTCACACCAAGACATTCAGGGGAAAAGCCTT	1860	
Qy	1774	ttatgtgcagggagtgtygcagaaggttttcggcaagaacctaactgtttaaagcaaga	1833	
Db	1861	TTATGTGACAGGAGTGTGGCAGAAAGTTTGGCAGAAAGCCTTAACCTGTTTAGCACAAGA	1920	
Qy	1834	gggcacactcagtgctcttgtgtgcagggagtgtyggcgaaggtcttgtctgaagttaa	1893	
Db	1921	GGGCACACTCAGAGTCTTTGTGTGCAGGAGTGTGGCAAGCTTTGTGTGAAGTTAA	1980	
Qy	1894	ctctcattaaacacagagagacacagcagggggagaagcctatgtgtcagggagtg	1953	
Db	1981	CTCTCATTAACACCAAGAGACACACAGGGGGAAGCCTATGTGTCCAGAGAGTGTG	2040	
Qy	1954	ggcaagcctttagccggcagtcacactcctatagaacacagaagacaattcaagagaga	2013	
Db	2041	GGCAAGGCTTTAGCCGGCAGTCCACACCTCATTAAGACACACAGACACTTCAGAGAGA	2100	
Qy	2014	agccttatattggcagaagatgtggagcgggtttagtcggaagtcgaaccttatcagag	2073	
Db	2101	AGCCTTATATTTTGCAGAAAGTGTGAGCGGGCTTTATGTGTGGAAGTCAAACTTATCCAG	2160	
Qy	2074	atcagaagacacactcagagatagaacttatgttataagggaatgtgtltagacacttla	2133	
Db	2161	ATCAAGGACACACACAGATTAAGAACTTTATGTGTATAGGAATGTGGTACAGCCTTA	2220	
Qy	2134	gccagagatcatalacttcatcagaacacagagaacacacagtcgtgtgttltcagc	2193	
Db	2221	GCCAGAGATCAATCACTTCAATCAGACACAGAGACACACACAGTGTGGCTTTTTCAGC	2280	
Qy	2194	catgtctagatatacgaatgtgagacatctgtgtgtgtttagttagtagagctgtactgt	2253	
Db	2281	CATGTCTAGATTAACAAAGTGGAGACATTTGTGTGTATGATGCAATGAGACTGTACTGT	2340	
Qy	2254	aagacttlatctccatccactcgaagagaatltgcgttcaatttcaaggagcccttgc	2313	
Db	2341	AAGACTTGTATCTCATCTCACCTGAAGGAAATCTGTGGTCAATTTTCAGGAGCCCTGCC	2400	
Qy	2314	cttctcaactgtgtagtgtgtgtgtgtgaaaccggtltagtgaatgtatgtgcagagag	2373	
Db	2401	CTTCTCACTGTGTGATGTGTGGTTGTGAAACCCCGTCAAGTAATGATGTGGCGAGAG	2460	
Qy	2374	cagtcaaatgcccaagcagatagaagggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2433	
Db	2461	CAGTAAATATGCCAGGACGATATGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2520	
Qy	2434	gacagttcccggttaaatcctcaactgaatgtgaagacgtgtcttccacttgtgtgtct	2493	
Db	2521	GACAGTCCGGGTAAATCTCTATACTGAATTGAGAACTGTGTCTTCCATTTGTGTGTGT	2580	

QY	2494	tcctcgaattatcccaacccctcaacatttacttactgactgaccttactaattggt	2553
Db	2581	TCCTCGATTGATGCCAACACCTTCACCTAATTTAGTATACCTGGCCCTTTCCTAATTTGT	2640
QY	2554	tttcaactgtctgtgtcccaaccttcttgagtggtgtcccttgataacttaacaaacagtcac	2613
Db	2641	TTTTACACTGTGGTGTGCCACACTTTTGAATGGTGCCCTTGCAATACTTACAATCACTCAAC	2700
QY	2614	gtgtattcccatctctgtgagcccaataaagaagccagactcagctgagtgagagagaaa	2673
Db	2701	GTGTATTCCTCCATATCTCAGCCCAATMAAAGACCCAGACTCAGCTGAGAGAGAGAAA	2760
QY	2674	tcacccctctctgtgaggtgttgaggaccactccctgataccctctccactgagagtgtc	2733
Db	2761	TCACCTCTCTGTGGGGGTGGGGACCACTCCCTGCATCCCTCTCACTGAGAGCTGTTC	2820
QY	2734	tttgcatacaaaattcttcttaccatctc	2767
Db	2821	TTTTGCTCAATMAAATTTCTTTTCTATCCATCTC	2854
RESULT	3		
LOCUS	AC008806/c	135173 bp	DNA linear PRI 27-APR-2000
DEFINITION	Homo sapiens chromosome 19 clone CTD-2086020, complete sequence.		
ACCESSION	AC008806		
VERSION	AC008806.4	GI:7656695	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 135173)		
JOURNAL	DOE Joint Genome Institute and Stanford Human Genome Center.		
REFERENCE	Direct Submission		
AUTHORS	Unpublished		
JOURNAL	2 (bases 1 to 135173)		
REFERENCE	DOE Joint Genome Institute.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE	3 (bases 1 to 135173)		
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (27-APR-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
AUTHORS	On Apr 27, 2000 this sequence version replaced gi:6600982.		
JOURNAL	Draft Sequence Produced by DOE Joint Genome Institute		
REFERENCE	www.jgi.doe.gov		
AUTHORS	Finishing Completed at Stanford Human Genome Center		
JOURNAL	www.shgc.stanford.edu		
REFERENCE	Quality: Phrap Quality >=40 99.5% of Sequence;		
AUTHORS	Estimated Total Number of Errors is 0.6.		
JOURNAL	SHGC-141611 G63592		
REFERENCE	SHGC-37349 G30531		
AUTHORS	SHGC-86030 G61627.		
JOURNAL	Location/Qualifiers		
REFERENCE	1..135173		
AUTHORS	/organism="Homo sapiens"		
JOURNAL	/db_xref="taxon:9606"		
REFERENCE	/chromosome="19"		
AUTHORS	/clone="CTD-2086020"		
JOURNAL	BASE COUNT 41590 a 27234 c 27240 g 39109 t		
ORIGIN			
QY	427	cagatcgaagcgcgaataactactatgctcctctgcctctgatttctccagtcagc	486
Db	126941	CAGATTCGAGGCGCAAAATTCATCTATGCTCCCTGCGCTCTGATTTTTCACACTGAGC	126882
Query Match	81.0%, Score 2244, DB 9, Length 135173;		
Best Local Similarity	99.9%, Pred. No. 0;		
Matches 2344, Conservative	0; Mismatches 2; Indels 0; Gaps 0;		

Oy	487	aagctctcagccacaatgctgctgtggtcatctctcagctgtttcaagtttatggg	546
Db	126881	AAGCTCTCAGCCAAACATGTGTGGCTGTAGTCTCTCTCTGACCTCTTTCAAGTTATGGG	1268222
Oy	547	caggaaatccctctcaccctgtggaaaacactctccagaagctcagaacaacagcagatc	606
Db	126821	CAGGAATCTCTCCACCCTGGGAAAACACTTCCAGAAATCAGAAAACACAGCAGATC	126762Z
Oy	607	catctgctttagtgygcanaagcagaatgaltccaagagggaggaagatccagactctgt	656
Db	126761	CATCTGCTTTAGTGGCAACAGAAATGGATTCAAGAGGGAGAAAGACTCCAGACTCTGT	1267020
Oy	667	ttggagagtaagcaaaaaatggaacttcaaggactttccagccacctgaagaacac	726
Db	126701	TTGGAGAGTAAAGCAAAAATGGCACTTCAAGGCACCTTCCAGCCCACTGAAGAACAC	1266421
Oy	727	agccagacaaatcccaagaagaacaacagtggtgatagagttcagagccctgaacgga	786
Db	126641	AGCCAGCACTCCAAAGAAACAAACAGTGGTATATAGGGTCCAGCCCTGAACGGA	1265822
Oy	787	ggcgcagatcctgaggaacaacagacaagtatgca tgg tttagaagttccagatlttgyag	846
Db	126581	GGCGAGATCTGAGAGAAACACAAAGATGTGCATGGTTTAAAGTGTTCAGAAATTGGAG	1265222
Oy	847	aaatcaaatatgaagaagtttggccagggcttaltcaagagtgcaaacctcttagctcc	906
Db	126521	AAATCAAAATATGAAGAGATTTGGCCAGGCTTTATCAAGAGAGTCAAACTCTTTAGCTTCC	126462Z
Oy	907	agaagacacaacacgtgggagagcaccttaacatgatactgtgtgggagagaagcttggca	966
Db	126461	AGAAACACCAACATCGGGAGACACTTTCATGTAACCTGATGGGGAGACAGCTTTGGCA	1264020
Oy	967	gtatgcagctccatcaaaaacccaagacacactctggyggaagaccttatgtgtgca	1026
Db	126401	GTATGCTAGTCTCTCATCAAAAAACCAAGACACACTCTGGGGAAACACTTATGTGTGCA	126342Z
Oy	1027	gggaatgtgggcgaagctttagctgtgaagltcaaacctgatacacatcagaagacacat	1086
Db	126341	GGGAATGTGGCGCGAGGCTTTACGTGGAAGTCAAACTGATCACACATCAAGAGCACACT	126282Z
Oy	1087	caggggagaaaccttatgtgtgcagaagtgttgaacgaagcttacttggaaatcgaaac	1146
Db	126281	CAGGGAGAAAACCTTATGTGTGCAAGGATTTGTGACACAGGCTTTACTTGGAAAGTCGAAC	126222Z
Oy	1147	tcttaacacatcagcggagacacaccgaaggtcacaagcttatgtgtgaagaagatgtgggc	1206
Db	126221	TCTTTACATCAGCGGACACACTCAGGCTCAAGGCTTATATGTGTCAAGGAATGTGGGC	126162Z
Oy	1207	agagctttagccttgaagtcaaacctcatlaccacacagagggcgcaacactvgggagaaagc	1266
Db	126161	AGAGCTTAGGCTGTAAGTCAAACTCTATTTACCACAGAGGGGCACTAGGGGAGAGGC	126102Z
Oy	1267	cttatgttctcagggaaatgtvggcgttgcagcatlccacacccgtgtcagaaca	1326
Db	126101	CTTATGTTTGCAGGGAATGTGGGCTGTGCTTCCGCAAGCATTCACACTGTGACAGACAA	126042Z
Oy	1327	agagagacacatcagagagaagaagccttaacatttcagggagtgtgagacaaggtctttagcc	1386
Db	126041	AGAGAGCACATTTCAAGGAGAGGCTTTACATTTCCAGGGAATGTGAACCAAGGCTTTAGCC	125982Z
Oy	1387	agaagtcacacctcaacgaacacttlaagacacacagagagaaagccttatgtgca	1446
Db	125981	AGAATTCACACTCATCAGACACTTAAAGACACACACAGAGAGAAACCTTATGTATGCA	125922Z
Oy	1447	cagaatgtgggcgttcaactttagctgtgaaatcaaacctcnaaacacccagaggaacacat	1506
Db	125921	CAGAATGTGGGCGTCACTTTAGCTGGAATATCAAACTCAAAAACACACAGAGCACACT	125862Z
Oy	1507	caggggtttaaacccttatgtctgcctctgaggtgcggcaggtgtcttagcctgaatgtcaaac	1566
Db	125861	CAGGGTTAAACCTTATGCTGCTGTGAGTCCGGGCAAGTGTGTTAGCTTAAGTCAAAAC	125802Z

QY	1567	tttaacaaacccagaggtgcacacagggggaggaagccaattgtaatgtaacagagtggtgggc	1526
Db	125801	tttaacaaacccagaggtgcacacacgggggaaagccatttggtaatgtaacagagtggtgggc	1257422
QY	1627	gaagccttaaccccggaatacaaccctgtaacagcaacagagagacacactcaagggagagc	1586
Db	125741	gagggcttttaaccccggaatacaaccctgtaacagcaacagagagacacactcaagggagagc	125682
QY	1687	caattgtatgtgcttgatgtgtggaagagccttaatgataagatccaccctcatccaacc	1746
Db	125681	caattgtatgtgcttgatgtgtggaagagccttaatgataagatccaccctcatccaacc	1256222
QY	1747	aaagagacacattcaaggggaaaaagcccttatgtgcaaggggtgtggagaagtttcggc	1806
Db	125621	aaagagacacattcaaggggaaaaagcccttttatgtgcacagggatgtggcgaaggtttcggc	1255622
QY	1807	agaagccctaaccctgtttaagcacagaaggacacactcaagtgccctgtgtgcaggagat	1866
Db	125561	agaagccctaaccctgttttaagcacagaaggacacactcaagtgccctgtgtgcaggagat	125502
QY	1867	gtgggcaagggcctttgtgtcctaagtaactcataltaaacacagagagacacgcagggg	1926
Db	125501	gtgggcaagggcctttgtgtcctaagtaactcattttaaaccacagagagacacgcagggg	125442
QY	1927	ggaagccctcatgtgtgcaggagaggtgtgggaagccttaacggcgagtcacacctatta	1986
Db	125441	ggaagccctcatgtgtgcaggagaggtgtgggaagccttttaacggcgagtcacacctatta	125382
QY	1987	gacacacagaagacacattcagaagagagaagccctatcttgacaagatgtgtgacgggct	2046
Db	125381	gacacacagaagacacattcagaagagagaagcccttattttgcagaagatgtgtgacgggct	125322
QY	2047	ttaagtcggaagtccaaccttaacagatacgaagagacacactaagatagaacttatgt	2106
Db	125321	tttagtcggaagatccaacctttatcagacatcagagagacacactaagatagaacttatgt	125262
QY	2107	tgtatagggaatgtgtgtaacagccttttagccagggatcaactcttaacagacaccaga	2166
Db	125261	tgtatagggaatgtgtgtaacagccttttagccagggatcaactcttaacagacaccaga	125202
QY	2167	cacacacagtgctgtggtccttttaagccattgtcctagaatacacaagatgtgaacatctgt	2226
Db	125201	cacacacagtgctgtggtccttttaagccattgtcctagaatacacaagatgtgaacatctgt	125142
QY	2227	tgtgattatgcataagagactgtactcgtgtaagacttgtatctcaatccaactgaagaagat	2286
Db	125141	tgtgattatgcataagagactgtactcgtgtaagacttgtatctcaatccaactgaagaagat	125082
QY	2287	tgtcgtggccattttcagaagaagccgtcccttcctccacgtgtgagtgtgggtgttggaaac	2346
Db	125081	tgtcgtggccattttcagaagaagccgtcccttcctccacgtgtgagtgtgggtgttggaaac	125022
QY	2347	cggtcaggttaatgtatgtgcaggaaggcaatcaaatgtcccaagcagatatagggtgtggtac	2406
Db	125021	cggtcaggttaatgtatgtgcaggaaggcaatcaaatgtcccaagcagatatagggtgtggtac	124862
QY	2407	ctgtgtgaaaceccaacttaaaagctgaaagacagtcctccggtctaaatctcatactgaaatga	2466
Db	124961	ctgtgtgaaaceccaacttttaaagctgaaagacagtcctccggtctaaatctcatactgaaatga	124902
QY	2467	gaacctgtcttcccaattgtgtgtgtcttctccgatgtatgcccaaacctccaactatltt	2526
Db	124901	gaacctgtcttcccaattgtgtgtgtcttctccgatgtatgcccaaacctccaactatltt	124842
QY	2527	acgtataacccgtcccttcccaatgtgtgtttacacgtcgtgtgccacacttttgagtggtg	2586
Db	124841	acgtataacccgtcccttcccaatgtgtgtgtttacacgtcgtgtgtgccacacttttgagtggtg	124782
QY	2587	cccttgatacttaacaaatcagtcaaaatgtatctccctattcttgagccataaagaagc	2646
Db	124781	cccttgatacttaacaaatcagtcaaaatgtatctccctattcttgagccataaagaagc	124722
QY	2647	cagaactcagctgcagtgtagagagaagaatcaaccctgtcgtgtgaagtttggggaccactccct	2706

```

Db 124721 CAGACTCAGCTGCAATGAGAGAAACACCCTCTGTGGGGTTGGGAGACACTCCCT 124662
OY 2707 gatccccctccatctgagagctgtcttttctcaataaaatctcttccaccatcc 2766
Db 124661 GCATCCCTCTCCATGAGAGACTGTCTTTTCTCAATAAATTTCTTTCACCATCT 124602
OY 2767 caacct 2772
Db 124601 CACCT 124596

RESULT 4
AC016590 253217 bp DNA linear HTG 20-APR-2001
LOCUS AC016590
DEFINITION Homo sapiens chromosome 19 clone CTD-3220F14, WORKING DRAFT
SEQUENCE 39 unordered pieces.
AC016590
AC016590.6 GI:13699590
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 253217)
TITLE DOE Joint Genome Institute.
JOURNAL Sequencing of Human Chromosome 19
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 253217)
TITLE DOE Joint Genome Institute.
JOURNAL Direct Submission
AUTHORS Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Apr 20, 2001 this sequence version replaced gi:7711568.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web Site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 940643, BC905667
Center clone name: C1MB-EL_3220F14
-----
Summary Statistics
Consensus quality: 207648 bases at least Q40
Consensus quality: 225569 bases at least Q30
Consensus quality: 232428 bases at least Q20
Estimated insert size: 250510; agarose-fp estimation
Estimated insert size: 249417; sum-of-contigs estimation
Quality coverage: 9.03 in Q20 bases; agarose-fp estimation
Quality coverage: 9.07 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'Working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1019: contig of 1019 bp in length
* 1020 1119: gap of unknown length
* 1120 2152: contig of 1033 bp in length
* 2153 2252: gap of unknown length
* 2253 3701: contig of 1449 bp in length
* 3702 3801: gap of unknown length
* 3802 5053: contig of 1252 bp in length
* 5054 5153: gap of unknown length
* 5154 6184: contig of 1031 bp in length
* 6185 6284: gap of unknown length
* 6285 7809: contig of 1525 bp in length
* 7810 7909: gap of unknown length
* 7910 8986: contig of 1077 bp in length
* 8987 9086: gap of unknown length

```

* 9087 10090: contig of 1004 bp in length
* 10091 10190: gap of unknown length
* 10191 11225: contig of 1035 bp in length
* 11226 11325: gap of unknown length
* 11326 11365: contig of 1040 bp in length
* 11366 12465: gap of unknown length
* 12466 14535: contig of 2070 bp in length
* 14536 14635: gap of unknown length
* 14636 16489: contig of 1854 bp in length
* 16490 19109: gap of unknown length
* 19110 19209: contig of 2520 bp in length
* 19210 21841: gap of unknown length
* 21842 21942: gap of unknown length
* 21943 23735: contig of 1794 bp in length
* 23736 25835: gap of unknown length
* 25836 25961: contig of 2026 bp in length
* 25962 29117: gap of unknown length
* 29118 31907: contig of 2690 bp in length
* 31908 32008: gap of unknown length
* 32009 36364: contig of 4357 bp in length
* 36365 43309: gap of unknown length
* 43310 43409: gap of unknown length
* 43410 48185: contig of 4776 bp in length
* 48186 48286: gap of unknown length
* 48287 55626: contig of 7341 bp in length
* 55627 55727: gap of unknown length
* 55728 65781: contig of 10055 bp in length
* 65782 65882: gap of unknown length
* 65883 69673: contig of 3792 bp in length
* 69674 79389: gap of unknown length
* 79390 79489: gap of unknown length
* 79490 91453: contig of 11964 bp in length
* 91454 91553: gap of unknown length
* 91554 96672: contig of 5118 bp in length
* 96673 100544: gap of unknown length
* 100545 100644: contig of 3773 bp in length
* 100645 111423: gap of unknown length
* 111424 111523: gap of unknown length
* 111524 115278: contig of 3755 bp in length
* 115279 115379: gap of unknown length
* 115380 124011: contig of 8633 bp in length
* 124012 124112: gap of unknown length
* 124113 138475: contig of 14364 bp in length
* 138476 138575: gap of unknown length
* 138576 140613: contig of 2038 bp in length
* 140614 140713: gap of unknown length
* 140714 144949: contig of 4236 bp in length
* 144950 145049: gap of unknown length
* 145050 164352: contig of 19303 bp in length
* 164353 164452: gap of unknown length
* 164453 170540: contig of 6088 bp in length
* 170541 170640: gap of unknown length
* 170641 185046: contig of 14406 bp in length
* 185047 185146: gap of unknown length
* 185147 222209: contig of 37063 bp in length
* 222210 223309: gap of unknown length
* 223310 253217: contig of 30908 bp in length.

Location/Qualifiers

1..253217

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="19"

/clone="CTD-3220F14"

/clone_lib="Caltech human BAC library D"

BASE COUNT 62246 a 62867 c 62532 g 61752 t 3820 others
ORIGIN

Query Match 81.0%; Score 2244; DB 2; Length 253217;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2344; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 427 cagaatcgaagccagaatcaactagtcctccctccctcgtatttccagtcagc 486
|||||
Db 74511 CAGAAATCGAAGCCAGAAATTCACATTAAGTCCCTCCCTGATTTTCACATGAGC 74570
OY 487 aagctctagcccaacatctgtgcgtgagtcacatctcccaatgctttcaagttatgg 546
|||||
Db 74571 AAGCTCTAGCCCAACATGTGGCTGAGTCAATCTCTCAGCTGTTTCAAGTTATGG 74630
OY 547 caggaaatccctccaccctgggaaacacataccagaagatcaagaacaacagcagatc 606
|||||
Db 74631 CAGGAATCCCTCCACCTGGGAAACACTATCCAGAAATCAGAAACACAGCAGATC 74690
OY 607 catctcgtcttagtgcgaagcagaaatgattcaagaaggagaaagactccagactctgt 666
|||||
Db 74691 CATCTGCTTTAGTGGCAAGCAGAAATGATTCAGAGGAGAAACATCCAGACTCTGT 74750
OY 667 ttggagagatgaagcaaaatggaacttaagaagcacttccagccaccctgaagaacac 726
|||||
Db 74751 TTGGAGAGATGAAGCAAAATGGCACTTCAAGGCACTTCCAGCCCTGAAGAACAC 74810
OY 727 agccagacagctcccaaggaagacaacacagctgtgtatagtggtccagccctgaagcga 786
|||||
Db 74811 AGCCAGCACAGTCCACAGGAAAGACACACAGTGTGATATAGGTCCAGCCCTGAAGC 74870
OY 787 gggcagatctaggaagaaacagacaaatgcatgattgaagctcaggaatttgag 846
|||||
Db 74871 GGGCAGATCTAGAGAAACAGCAAAATGATTCAGTGTGATTAAGAACTCAGGATTTGAG 74930
OY 847 aaatcaaatatgaagaagtttggccagcttatacaaggaatgcaactcctttagctcc 906
|||||
Db 74931 AAATCAATATGAAGAGTTTGGCCAGGCTTATCAAGAGCAAACTCTTACCTCC 74990
OY 907 agaagacaacactgggagagacacttcatgtacatgagtgaggagagacacttggca 966
|||||
Db 74991 AGAAGACACAAACTGGGAGAGACACTTACATATACATGAGTGGGAGACACTTTGGCA 75050
OY 967 gtaatgcagctccatcaaaaaacccaagagacacactctgggggaaagccttatgtgtca 1026
|||||
Db 75051 GTATGTCACTCCTCATCAAAAAACCCAGGACACACTTGGGGGAAAGCCTTATGTGCA 75110
OY 1027 ggggaatgtgggagagccttccgttggaagtcaaacctgatacaataagagaaacact 1086
|||||
Db 75111 GGAATGTGGGGAGGCTTTAGGTGAAGTCAAACTGATCAACATCAGAGACACACT 75170
OY 1087 cagggagaacacttatgtgtcaaggaatgtgtgacgaagccttacttggaaagtcgaac 1146
|||||
Db 75171 CAGGGAGAAACCTTATGTGTGTCAGAGATGTGGAGAGAGGCTTACTTGGAAAGTGGAAC 75230
OY 1147 tctttacacatcagcgcgacacacatcaaggcctcaatgtgtgtgcaaggaatgtgggc 1206
|||||
Db 75231 TCTTTACACATCAGCGGACACACTCAGGCTCAACCTTATGTGTGCAAGGAATGTGGGC 75290
OY 1207 agagctttagcctgtgaatgtcaaacctcatcacccagagagcgacacttgggggaagc 1266
|||||
Db 75291 AGAGCTTTAGCTGTAAGTCAAACTCATTTACCCACAGGAGGCGACATCGGGAGAAC 75350
OY 1267 cttaatgttcaggaagatgtggcgttgcgcagacatcaacaactgtgtcgaacaca 1326
|||||
Db 75351 CTTATGTTTGCAGGAAATGTGGCGTGTGCGCAGCATTCACACCTGTCAGACACACA 75410
OY 1327 agagacacatcagagagagaagccttaccattgtcaggaagtgtgaagagccttaagc 1386
|||||
Db 75411 AGAGACACATTCAGAGAGAGAGCCTTACATTTGAGGAGAGTGTAGCAGACGCTTATAGC 75470
OY 1387 agaagtcacactcactgaagacacttaagagacacacagaggaagagccttaagtatga 1446
|||||
Db 75471 AGAAGTCAACCTCATCAGACACTTAAGAGACACACAGAGAGAAAGCCTTATGATGCA 75530

QY	1447	caaatgltggtgcacctttagctgtgaaatcaaaccttaaaacacccagagacacact	1506
Db	75531	CAGAATGTGGGGCTCACTTTAGCTGGAAATCAAACTCAAAACACACGAGAGACACT	75590
QY	1507	caggggttaaacccttctgtcgcctgagtgccggcagtgctttagccgtgaatcaaac	1566
Db	75591	CAGGGGTTAAACCTTATGCTGCTGGATGCGGGCAGTGCTTTACCCTGAACTCAACC	76505
QY	1567	ttaaacaacaccaagagtcacacacgggggaaagaaaccatttgtatagagtgtyggc	1626
Db	75651	TTTAACTAAACACAGAGTGCACACACGGGGAGAAAGCATTTGTATAGGAGTGTGGGC	75711
QY	1627	ggggcttcccggaatacaaccctgatacgcacagagagacacctaagggaaagc	1686
Db	75711	GGGGCTTACCCGGAATATACCCCTGAGACGCACGAGAGACACTAGGGGAACAC	75770
QY	1687	catctgtatgtctgagtgctggacagagctttaatgatatccacccttaacac	1746
Db	75771	CATTTGTATGTCTGAGTGTGGACAGAGCTTTATATATAAGTCCACCCTCATTTCCACC	75833
QY	1747	agaagacacatcagggggaagaccttltatgtccaggagtgtygcagaagtltcgcg	1806
Db	75831	AGAGGACACATTTACAGGGGAAAAAGCCTTTATGTGACAGGAGTGTGCACAAAGTTTCGGC	75890
QY	1807	agaagcctaacctgttttagcacaagagagacacaccagtgcccttgttgcagagat	1866
Db	75891	AAAGCCTTAACCTGTTTATAGGCACAAAGGGCACACACTCAGTGTGCTGTGACAGGAGT	75950
QY	1867	gtcggcacaagccttctgtcctaagctaacctcaataaacacagagagacacagcaggg	1926
Db	75951	GTGGCGAAGGCTTTTGTCTGATGATTACCTCATTTAAACACGAGAGCACACGACGGGG	76010
QY	1927	ggaagcctcaatgtgtgcagggagtgtyggcagaagctttagccgycagtcacacctcatla	1986
Db	76011	GGAAGCCTCATGTGTGACAGGAGTGTGGCGCAAGGCTTTAGCCGGCAGTCAACCTCATTA	76070
QY	1987	ggacacacagagacacatttcagagagaaacccctataatttgcagaaagtgtgacggggt	2046
Db	76071	GACACACAGGACACCTTCAGAGAGAAACCTTTATTTTGCGAAAGTGTGACGGGGCT	76130
QY	2047	tttagtcggaagtcacaacctcacaacacagagacacacccagatagaacctatgt	2106
Db	76131	TTTAGTCGGAAGTCCAACTTATCAGACATCAGAGGACACACTCAGATAGAAACTTTATG	76190
QY	2107	tgtataaggaaatgtgtacagacctttagccagagatcatalacttcacagacaccaaaga	2166
Db	76191	TGTATAGGAAATGTGTATACAGCCTTTAGCCAGAGAGTCATATTCATCAGACACCAAGAGA	76250
QY	2167	cacacacagtgctgtgtgaccttltcagccatgtctagataccaaagtgtgaacatctgt	2226
Db	76251	CACACACAGTGTGTGGCTTTTTCAGCCATTCCTGATACCAAAGTGAGACATTCCTGTG	76310
QY	2227	tgtgtatagcagtgagacatgtactcgttgaagctttrataccatccacactgaagagaa	2286
Db	76311	TGTGATATGCTATGAGACTGTACTGTGTAAGACTTGTATCTCATCCACTGAAAGGAAT	76370
QY	2287	tgtcgtcctcaatttcaagagccctgccttccctcaactgtgagatgtgtgtgtgtgaaac	2346
Db	76371	TGCTGGCTCATTTTTCAGGAGCCCTGCTCCTCACTGATGATGTGGGTGTGGAAAC	76430
QY	2347	cggtcagtgtaatgtatgtgcagagagcagtcacaaatgtcccagcagatagggtyggta	2406
Db	76431	CGGTACAGTAAATGATGTGGCGAGGAGGCGAGTCAAAATGCCACAGGAGATGTGGGTGTAC	76490
QY	2407	ctgtgtgaaacccaaccttaagccttaagacagatcccggttaatactcaatagtatgtga	2466
Db	76491	CTGTGTGAAACCCAACTTAAACCTGAAGACATCTCCGGCTTAATCTCTACTGAAATGA	76550
QY	2467	gaacctgtcttcccaattgtgtgtgtcttccctcgatgtgataccaaaccttccactatct	2526
Db	76551	GAACTGTCTTCCCATTTGGTGTGCTTCTCCGATGTATGCCAACCTTCACACTATTTT	76610
QY	2527	acgtataccgtcccttccctaattgtgttttaacactgtctgtgccacacttttagatgtgt	2586

Accession	Sequence	Position
Db	76611 AGGTAACTGCGCCCTTTTCCCTAATGTGGTTTTCACGTGCGTGCACCGCTTTTGAATGGG	76670
Qy	2587 ccttggcaacttaaaaaatcagtcacagcgtgtgtctcccatctcagaacccataaaagcc	2646
Db	76671 CCTTTGCATCTTACCAATCAATCACTCAACGTAATCCCTATTCTGACCCATAAAGACC	76730
Qy	2647 cgaactcagctgcagtlgagagagaataatcaccctgcctgtagaggttggagaccatccct	2706
Db	76731 CAGACTCAGCTCAGTGAAGAGAAATCACCTCTGCTGGGGGTTGGGAGACACTCCCT	76790
Qy	2707 gatcccccctcactgagagcgtgtcttcttgcataaaaaatctttctacccatcc	2766
Db	76791 GCATCCCTCTCTCACTGAGAGGTGTCTTTTCTCAATTAATTCCTTTTCTATCCATCCT	76850
Qy	2767 caacct 2772	
Db	76851 CACCTT 76856	

	RESULT	N	5	
Locus	AC093235	38173 bp	DNA	linear PRI 29-SEP-2001
DEFINITION	Homo sapiens chromosome 19 clone LNLNR-267A4, complete sequence.			
ACCESSION	AC093235			
VERSION	AC093235.2	GI:15808549		
KEYWORDS	HTG;			
SOURCE	human.			
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 38173)			
REFERENCE	DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Unpublished 2 (bases 1 to 38173) DOE Joint Genome Institute.			
AUTHORS	Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA 3 (bases 1 to 38173)			
JOURNAL	DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Submitted (29-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA On Sep 29, 2001 this sequence version replaced gi:15193369. Draft Sequence Produced by DOE Joint Genome Institute			
REFERENCE	www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www.shgc.stanford.edu Quality: Phrap Quality >-40 99.9% of Sequence; Estimated Total Number of Errors Is 0.			
AUTHORS	Location/Qualifiers 1..38173 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="19" /cclone="LNLNR-267A4"			
TITLE				
COMMENT				
FEATURES				
Source				
BASE COUNT	9495 a 8465 c 8118 g 12095 t			
ORIGIN				
Query Match	79.1%; Score 2193; DB 9; Length 38173;			
Best Local Similarity	99.9%; Pred. NO. 0;			
Matches 2343; Conservative	0; Mismatches 3; Indels 0; Gaps 0;			
OY	427 cagaatcgaagccaggaaattcaacttagtccttccgaccttgattttctcacgacg 486			
Db	28888 CAGATTCGAAGCACAATTCAACTTAGTCCCTCGCTGTGATTTTCACATGCAGC 28947			
Oy	487 aaagtccaaccacaatatgtggcgtaaatcatctctctcacgactgtttttaagttatgg 546			
Db	28948 AAGCCTTGAGCCAACATAATGGCGTAATCATCTCTCACGCTGTTTTCAAGTTAAATGSG 29007			

QY 547 caagaaatccctccacactggaagaacataccagaagatcagaagaacaacagagatc 606
|||||
Db 29008 CAGAAATTCCTCTCCACTGGGAAACACTATCCAGAAAGATCAGAAACAAACAGCAGGATC 29067
QY 607 catctgctttagtgcgaagcagaatgatlccaagagaggaagactccaagactcgt 666
|||||
Db 29068 CATTCCTCTTAGTGGCAAGCAGAAATGSAATTCAGAGGGAGAGACTCCGAGACTCCGT 29127
QY 667 ttggagagtaagcaaaaatggaactccaagcacttccaagcccaactgaagaac 726
|||||
Db 29128 TTGGGAGAGTAAGCAAAAATGGCACTTCAAAAGGCATTTCCAGCCCTCCTGAAAGAACAC 29187
QY 727 agccagacagctccaagaagaacacacacagtgatgatataggtccaagccctgaagga 786
|||||
Db 29188 AGCCAGACAGTCCAAAGAGAACAAACAGTGTGATATAGGGTCCAGCCCTTAACGGGA 29247
QY 787 gggcagatctagagaaacagacaaaglatgcatggttlaagaagctcaagatctgag 846
GGGAGATCTAGAGAAACAGACAAATGTCATGGTTTAGAGTCCAGAGATTTGGAG 29307-
Db 29248
QY 847 aaatcaaatatgaagattggccagcttatacaagagatcaaacctcctagctcc 906
|||||
Db 29308 AAATCAAAATATGAAGAGTTTGGGCCAGCTTTATCAAGGAGTCAAACTCCTTAGCCTCC 29367
QY 907 agaagacacaaactggggagagacacttactgatactgagtgaggagagacttgga 966
|||||
Db 29368 AGAAGACACAACTGGGGAGACACCTTACATGACCTGAGTGGGAGACAGCTTTGGCA 29427
QY 967 gtagtgcagctccatacaaaaacccaagacacactcggggagaaagccttatgtgca 1026
|||||
Db 29428 GTATGTAGTCTCATCAAAAACCCAGACACACTCTGGGGGAAAGCTTATGTGTCA 29487
QY 1027 gggaaatgvggagagctttagtgcgaagtcacaacctgatacaatcagaagacact 1086
|||||
Db 29488 GGGATATGTGGCGAGGCTTTACGTGGAAGTCMAACCTGATCACATCAGAGGACACACT 29547
QY 1087 caggagagaaacctatgtgtgcagaagatgtgcagagagcttacttgaagtcgaacc 1146
|||||
Db 29548 CAGGGAGAAACCTTATGTGTGCAGAAAGATGTGACAGGGCTTACTTGGAAAGTCGAACC 29607
QY- 1147 tcttaacacatcagcggagacactcaagagctcaagccttatgtgcagaagaaatgtggc 1206
|||||
Db 29608 TCTTTACATCAGCAGGACACACTCAGAGGCTTATGTGTGCAAGAAATGTGGGC 29667
QY 1207 agagctttagcctgaagtccaacctcattaccacagagggcgcaactgvggagaagc 1266
|||||
Db 29668 AGAGCTTTAGCCTGAAGTCAAACTCATTTACCCACAGAGGGCGCACACTGGGAGGAAGC 29727
QY 1267 ctatgtttgcagggaatgtggtgtgcttgcagacatccaactgtgtcagaaca 1326
|||||
Db 29728 CTATGTGTTTGCAGGGAATGTGGGCTGGCTTCCGACGATTCACACTGGTCAAGACCA 29787
QY 1327 agaagacacatcagagagaagccttatacttgcaggaagtgtgagcaagagctttagc 1386
|||||
Db 29788 AGAGGACACATTTAGGAGAGAACTTATGAGAGAGTGTGAGCAAGGCTTTAGCC 29847
QY 1387 agaagtaacacctatacagaacttaagagacacacacagagagaagagccttatgtatgca 1446
|||||
Db 29848 AGAAGTACACCTCATAGACACTTAAGGACACACAGGAGAAAGCTTATGATATCA 29907
QY 1447 cagaatgtggcgctcaacttgcgtggaatcaaacctcaaacacacagagagagacact 1506
|||||
Db 29908 CAGAAATGGGCACTCATTTAGCTGGAATCAAACTCAAAACACACAGAGGACACACT 29967
QY 1507 cagggttaaaccttatagtctgcctgagatgcgggagagtgctttagcctgaagtaaac 1566
|||||
Db 29968 CAGGGGTTAAACTTATGTGTGCTGGAGTGGGAGTGGCTTAGCTTGAAGTCAAAAC 30027
QY 1567 ttaacaacacacagagtcacacacgggggagagagccattgtatgtcaggaagtgtggc 1626
|||||
Db 30028 TTAACAACACACAGAGGTCAACACAGGGGAGAGCCATTTGTATGTACGAGAGTGTGGGC 30087

QY 1627 gaggcttaccgggaatatcaaccctgatcacgcacacagagacacactcagggagaagc 1686
|||||
Db 30088 GAGGCTTTACCCGGAAATCAACCTTGAGCACGCACAGAGACACACTCAGGGAGAAAGC 30147
QY 1687 catltgatgtctgagtgatgagcagagcttataatgataatgcaaccttataccacc 1746
|||||
Db 30148 CATTTGATGTGTGAGTGTGAGAGAGGCTTTAATGATGAATCCACCTCATTTACACAC 30207
QY 1747 agaagacacatcaggggaaagcccttatagtgcaggaagtgtgcagaagtgctggc 1806
|||||
Db 30208 AGAGGACACATTTAGGGGAAAAAGCCTTTATGTGCAGGAGTGTGGCGAAGGTTTGGCG 30267
QY 1807 agaagcccaacctttagagacaagaagagacactcaagtgctttagtgagagaggt 1866
|||||
Db 30268 AGAAGCTTACCTGTTTAGGCACAAAGAGGCACCTCAGGTGCTTGTGTGACAGGAGT 30327
QY 1867 gtggcagaagctttagtgcatactccatataacacacaaagagacacagcaggg 1926
|||||
Db 30328 GTGGGCAAGGCTTTGTGCTTAAGTTAACTTCATTTAAACACAGACACAGCAGAGGG 30387
QY 1927 ggaagctcaatgtgtgcaggaagtgtggaagcctttagccggagatcaactatla 1986
|||||
Db 30388 GGAAGCCTCATGTGTGCAGGAGTGTGGCAAGGCTTTAGCGCGCAGTCAACCTCATTA 30447
QY 1987 gacacagagacacacatcagagagaagccttatttgcagaagaagtgtgaagcggct 2046
|||||
Db 30448 GACACACAGAGACACATTCAGAGAGAAAGCTTATTTGAGAAAGTGTGAGGGGCT 30507
QY 2047 ttagtcggaagttccaaccttatacagacatcagaagacacatcagatagaacttatg 2106
|||||
Db 30508 TTAGTCGGAAGTCCAACTTATCAGACATCAGAGACACACTCAGAGATGAAGCTTTATG 30567
QY 2107 tgtataggaagtgtgtacagacactttagccagaggtactactatcagaacacagaga 2166
|||||
Db 30568 TGTATAGGAATGTGTGACACACCTTTAGCCAGGAGTCTTACTTATCAGAACACAGAGA 30627
QY 2167 cacacagatgctgtgctttagcatttcaagcattgcatatacaaaagtgtgaagacattgtg 2226
|||||
Db 30628 CACACACAGTGTGTGGCTTTTCAAGCCATTGCTGATGATACCAAGTGSAGACATTTGTG 30687
QY 2227 tgtatatacagatagactgtactgtlaagactgtatctccatccactcgaagaagat 2286
|||||
Db 30688 TGTATATGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 30747
QY 2287 tgcgtgctcatttccagagccctgccttccatcactgtgagtgtgtgtgaaacc 2346
|||||
Db 30748 TGCTGGCTCATTTTCAGAGGCCCTGCTTCTCCTGATGATGATGATGATGATGATGATGAT 30807
QY 2347 cgttcaagtaatgtagtgcagagagagagtcacaaatgcccagagcagatagaggtgtgac 2406
|||||
Db 30808 CGGTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 30867
QY 2407 ctgtgaaacccaaccttaagcttgaagacagtcgccggtcctaattccataactgtaatga 2466
|||||
Db 30868 CTGGGAAACCAACCTTAAAGCTGAGACAGTCCCGGCTAAATCCTCATCTGATTTGA 30927
QY 2467 gaactgtcttccaatgt 2526
|||||
Db 30928 GAACCTGTCTTCCATTTGGT 30987
QY 2527 agctataccctcccttccatattgttttatacactgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2586
|||||
Db 30988 AGGTATACCTGCTTCTTCTTATTTGTTTAACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 31047
QY 2587 ccttgcatactataaatacagtcagtgatattccctatcttgaagccataaagaagc 2646
|||||
Db 31048 CTTTGCACTACTTCAAAATCAGTCAACGTGATTCCTTATTCAGGCTTAAAGAACAC 31107
QY 2647 cagactgcagcaggtgagagagaataacacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2706
|||||
Db 31108 CAGACTCAGTGCAGTGTGAGAGAAATCACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 31167
QY 2707 gcatcccttccacatgagagctgtcttcttgcataaataatcttcttaccatcct 2766

|||||
Db 31168 GCATCCCTCTCCACAGAGCTGTTCTTCTCAATAAATTTCTTTCTACCCACTCT 31227
Qy 2767 caccct 2772
|||||
Db 31228 CACCTCT 31233
RESULT 6
BC004513 1998 bp mRNA linear PRI 12-JUL-2001
LOCUS BC004513
DEFINITION Homo sapiens, clone IMAGE:3928207, mRNA, partial cds.
ACCESSION BC004513
VERSION BC004513.1 GI:3325426
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LTLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amandastystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LTLNL at: http://image.llnl.gov
Series: IRAL Plate: 14 Row: b Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
FEATURES
source Location/Qualifiers
1..1998
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3928207"
/tissue_type="Ovary, adenocarcinoma"
/clone_id="NIH_MGC_9"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1..1280
/codon_start=3
/product="Unknown (protein for IMAGE:3928207)"
/protein_id="AAH04513.1"
/db_xref="GI:3325427"
/translation="HEYEVSGFEIKYEEFPGFIKESNLTSLQKTGTGPPVYTEM
GDSFGMSVLTKNPRTHSGKRPYVCEGREGFTWKSNIITHORHSGKRPVCKDCGR
GFTWKSNIITHORHSGKRPYVCEGREGFTWKSNIITHORHSGKRPVCKDCGR
RQSHLVRHRTKTHSGKRPYVCEGREGFTWKSNIITHORHSGKRPVCKDCGR
KSNLKHRTKTHSGKRPYVCEGREGFTWKSNIITHORHSGKRPVCKDCGR
TLTHORHSGKRPYVCEGREGFTWKSNIITHORHSGKRPVCKDCGR
FRHRAHSGKRPYVCEGREGFTWKSNIITHORHSGKRPVCKDCGR
RTHSGKRPYVCEGREGFTWKSNIITHORHSGKRPVCKDCGR"
BASE COUNT 575 a 458 c 510 g 455 t
ORIGIN
Query Match 63.9%; Score 1771; DB 9; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1941; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 826 tagaagctcagaagatttgagaagaatcaaatatgaagagtttggccaagcttatacaag 885
|||||
Db 10 TAGAAGTCTCAGGATTTGGAGAATCAATATGAGAGTTTGGCCAGGCTTTATCAAG 69
Qy 886 agtcaaacctccttagctctcagaagaacaacatgagggaacacctatcatgtacacty 945
|||||
Db 70 AGTCAAACTCTTGTAGCCCTCCAGAAAGACAAACTGGGAGACACTTTCATGTACACTG 129
Qy 946 agtgggaagaagctttgacagtatgtacgtctcaatcaaaaacccaagaacacactctg 1005
|||||
Db 130 AGTGGGAGACAGCTTTGGCAGTATGTCACTCTCAATAAACCAGACACACTCTG 189
Qy 1006 ggggaagaagcttatgttgcagaagaatgaggcagagctttagcttgaagatcaaacctga 1065
|||||
Db 190 GGGGAAGAGCTTATGTGTCCAGGAAATGTGGGAGAGCTTTACGTGGAAGTCAAACTGA 249
Qy 1066 tcacacatcagaagaacacactcaagggaagaacacttatgttgcagaagattgtgcagag 1125
|||||
Db 250 TCACACATCAGAGACACACTCAGGGGAGAAACCTTATGTGCAAGATGTGAGAGAG 309
Qy 1126 gcttacttggaagtcgaacacctcttaacacatcaggaacacactcagaaggtcagaagctt 1185
|||||
Db 310 GCTTACTTGGAGTGAACCTCTTTACATCAGGAGACACACTCAGGCTCAAGCTT 369
Qy 1186 atgtgtcagaagaatggtgcagaagctttagcttgaagatcaaacctataccaacaga 1245
|||||
Db 370 ATGTGTGCAAGGAATGTGGGAGAGACTTATAGCTGAGTCAAACTATACCAACAGAGA 429
Qy 1246 gggcgaacatgagggaagaagccttatgttgcagaagaatggtgcgttgcagcagc 1305
|||||
Db 430 GGGCCACACACTGGGAGAGAGCTTATGTGAGGAGAAATGTGGGCTTCCACAGC 489
Qy 1306 attcacactggtcagaagaagaagaacacatcaggaagaagccttaccatttgcagag 1365
|||||
Db 490 ATTCAACCTGTGTCAGACACAAAGAGACACATTCAGGAGAGAGCTTATTCAGAGG 549
Qy 1366 agtgtgacaaagctttagccagaagtcacacactcaacacacacttaaggagacacacag 1425
|||||
Db 550 AGTGTGACGAAGGCTTTAGCCAGAAAGTCAACACTCATCAGACACTTAAGGAGACACAG 609
Qy 1426 gagagaagccttatgtatgcacagaaatggtgcagcactttagcttgcagaatcaaacctca 1485
|||||
Db 610 GAGAGAAGCTTATGTATGACAGAAATGTGGGCTGTACCTTATGTGTAATCAAACTCA 669
Qy 1486 aaacacacacagaagaacacactcaagggttaaaccttatgtcgtccttgaagtcaggagcag 1545
|||||
Db 670 AAACACACACAGAGACACACTCAGGGGTTAAACCTTATGTGTGCTGAGTGGGCAAGT 729
Qy 1546 gctttagcctgaagtcacacacttcaacacacacacagaagtcacacacgagggaagaagcag 1605
|||||
Db 730 GCTTATAGCTGAAAGTCAAACTTAAACAAACACAGAGTCAACAGGGGAGAGCCAT 789
Qy 1606 ttgtatgtacagagatgtggtgcagagccttaccgggaatcaacacctgatacagcagca 1665
|||||
Db 790 TTGTATGTACAGAGTGTGGGCGAGGCTTTACCCGGAATCAACCTGTATCAAGCACCGA 849
Qy 1666 ggaacacacacagaagaagaacatttgaatgtgcgaatgtggaagaagccttataatgata 1725
|||||
Db 850 GGACACACTCAGGGAGAGAGCCATTGTATGTGTGAGTGTGAGAGAGGCTTTAATGTATA 909
Qy 1726 agtcaaacctcattcaacacacagaagaacacatcaggggaagaagcctttagtgcagag 1785
|||||
Db 910 AGTCACCTCATTTCAACACAGAGAGACACANTCAGGGGAAAGCCTTTATGTGTCAAGG 969
Qy 1786 agtgtgcagaagatgttgcagcaagagccttaacctgtttaggcacaagaaggcacaactag 1845
|||||
Db 970 AGTGTGCAGAGAGTGTGGCAGAGAGCCTTAACCTGTATAGGCACAAAGGCGACACTCAG 1029
Qy 1846 gtgccttgtgtgcagaagaagtggtgcagaagcctttagtgcagaagtaacttcataaac 1905
|||||
Db 1030 GTGCCCTTGTGTGTCAGAGAGTGTGGCAAGGCTTTGTGTGAAGTAACTCTCATTAAC 1089

QY 1906 accagagacacacgcaaggaggaagcctcgtatgtcaggaagatgtgcaagcctta 1965
 |||||||
 Db 1090 AOCNAGAGCAGACCCAGGGGGAGCCTCATGTGTGACAGGAGGTGGGACAGCTTTA 1149
 QY 1966 gccggcagtcacacccctatagacacacagagacacatcagagagaagccttattt 2025
 |||||||
 Db 1150 GCCGCGACTCAGACCTCATTTAGACACACAGAGACACATTCAGAGAGAAGCCTTATATT 1209
 QY 2026 gcaagaagtgtggaagggccttagtcggaagtcacacccctatagacatagagacac 2085
 |||||||
 Db 1210 GCAGAAATGTGTGACGGGGCTTTAGTGGAAAGTCCACCTTATTCAGACATCAGAGACAC 1269
 QY 2086 actcagagatagaacattatgtatgtataggaatgtgtacagcctttagccagagatcat 2145
 |||||||
 Db 1270 ACTCAGATAGAACTTATGTATGTATAGGAATGTGTACAGCCTTTAGCCAGAGATCAT 1329
 QY 2146 actcatalcagacacacagagagacacacagatgtgtgacccctttagccatgttagata 2205
 |||||||
 Db 1330 ACTTCATCAGACACACAGAGACACACACAGTGTGTGCTTTTTCAGCCATTGTAGATA 1389
 QY 2206 ccaaggtgagacattctgtgtgtatgatagatgagactgtactgtgtgaagcctgtatc 2265
 |||||||
 Db 1390 CCAAGGTGAGACATTTCTGTGTGTATGTATGTATGTATGTATGTATGTATGTATGTAT 1449
 QY 2266 tccatccacccctgaagagaatgtgtgtgtcattcaggaagccctgtcctcactgtt 2325
 |||||||
 Db 1450 TCCATCCACCTGAAAGAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1509
 QY 2326 ggaatgt 2385
 |||||||
 Db 1510 GGATGT 1568
 QY 2386 caggaagatagaaggt 2445
 |||||||
 Db 1569 CAGGACGATAGGGGT 1628
 QY 2446 taatcccatcactgaatgtgaagacctgtctcctcattgtgtgtgtgtgtgtgtgtgtgt 2505
 |||||||
 Db 1629 TAAATCCGATATCTGAATGTGAAGAACCTGTCTCCCATTTGGTGTGTGTGTGTGTGTGT 1688
 QY 2506 tcccaaccctcactatatttaagatatacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2565
 |||||||
 Db 1689 TCCCAAACCTTCACCTATTTTACGTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1748
 QY 2566 gtgtccaccccttgaatgt 2625
 |||||||
 Db 1749 GTGCCACCTTTTGT 1808
 QY 2626 attctgaagccataaaagacccagacatcagctgtcagtgagagagaatcacccctgtgt 2685
 |||||||
 Db 1809 ATTCTGAGCCCATTAATAAGACCCAGACTCAGTGTGAGTGTGAGAGAAATCACCTGTGTGT 1868
 QY 2686 ggaagtttgggagacacccctcgtcgtatccctcctcagctgtgtgtgtgtgtgtgtgtgt 2745
 |||||||
 Db 1869 GGGGGTTGGGGACACTCCCTGACATCCCTGTCCACTGTGAGACTGTCTTTTGTGCTCAATA 1928
 QY 2746 aaattcttctacccatccca 2768
 |||||||
 Db 1929 AAATTCTTTTCTACCCATCTCA 1951

RESULT 7
 HUMHKL
 LOCUS HUMHKL 672 bp DNA linear PRI 27-APR-1993
 DEFINITION Human Kruppel related gene, exon X, clone pHKRLS.
 ACCESSION M20675 M19429
 VERSION M20675.1 GI:184108
 KEYWORDS Krueppel-related protein.
 SOURCE Human DNA, clone pHKRLS.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 672)

AUTHORS Ruppert, J. M., Kinzler, K. W., Wong, A. J., Bigner, S. H., Kao, F. -T.,
 Law, M. L., Senear, H. N., O'Brien, S. J., and Vogelstein, B.
 TITLE The Gli-Kruppel family of human genes
 JOURNAL Mol. Cell. Biol. 8, 3104-3113 (1988)
 MEDLINE 89096896
 COMMENT computer-readable sequence for [1] kindly provided by J. M. Ruppert,
 20-JUN-1988.
 FEATURES
 source location/Qualifiers
 1..672
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 exon
 <1..>672
 /note="exon X"
 BASE COUNT 195 a 155 c 184 g 138 t
 ORIGIN Chromosome 19.

Query Match 24.2%; Score 672; DB 9; Length 672;
 Best Local Similarity 100.0%; Pzed. No. 0;
 Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 942 actgagtgaggagagacgtttgtgcagtgatgtatgtccatcaaaaacccaagagacac 1001
 |||||||
 Db 1 ACTGAGTGGGAGACAGCTTTGGCAGTATGTATGTATGTATGTATGTATGTATGTATGTAT 60
 QY 1002 tctgggggaaagccttatgtgtgtcaggaagtgtggcagagcctttagctgtgaagcaaac 1061
 |||||||
 Db 61 TCTGGGGGAAAGCCTTATGT 120
 QY 1062 ctgatacacatcagagagac 1121
 |||||||
 Db 121 CTGATCACACATCAGAGAGAC 180
 QY 1122 cgaagccttactgtgaatgtcgaacctttaaatacagcgaacacacacacacacacacacac 1181
 |||||||
 Db 181 CGAGCCTTACTTGAAGTCGAACTCTTTTACACATCAGCAGACACACACACACACACACAC 240
 QY 1182 ccttatgtgtcagaagt 1241
 |||||||
 Db 241 CCTTATGTGTGAGAGAAATGT 300
 QY 1242 cagaagggcgacacctgtgggagaagccttatgtttgcaaggaaatgtgtgtgtgtgtgtgt 1301
 |||||||
 Db 301 CAGAGGGCGCACACTGGGGGAGAGCCCTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
 QY 1302 cagcatcaccctgt 1361
 |||||||
 Db 361 CAGCATTCACACCTGT 420
 QY 1362 agggagtgtgcaagcctttagcagaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1421
 |||||||
 Db 421 AGGAGGTGTGACAGCAGCCTTTAGCCAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
 QY 1422 acagagagagaagccttatgtatgcagagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1481
 |||||||
 Db 481 ACAGGAGAGAAACCTTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGT 540
 QY 1482 ctcaaaacacacagagagac 1541
 |||||||
 Db 541 CTCAAAACACACAGAGAGAC 600
 QY 1542 cagtgctttagcctgagacaaaccttaacaaacacagaggttcacacacgggggaggaag 1601
 |||||||
 Db 601 CAGTGCTTATAGCCTGAAGCAACCTTAAACAAACACACAGAGTGTGTGTGTGTGTGTGTGTGT 660
 QY 1602 ccatttgtatgt 1613
 |||||||
 Db 661 CCATTGTATGT 672

RESULT 8
 LOCUS AX067361 678 bp DNA linear PAT 24-JAN-2001

DEFINITION	Sequence 65 from Patent WO0078960.
ACCESSION	AX067361
VERSION	AX067361.1 GI:12544985
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 678) Yugui,J. and Mitcham,J.L. Compositions and methods for the therapy and diagnosis of breast cancer
AUTHORS	
TITLE	
JOURNAL	Patent: WO 0078960-A 65 28-DEC-2000;
FEATURES	CORIXA CORPORATION (US)
source	Location/Qualifiers
' misc-feature	/organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	1..678 /note="n = A,T,C or G"
ORIGIN	165 a 167 c 160 g 182 t 4 others
?	
Query Match	20.6%; Score 571; DB 6; Length 678;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 571; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
Oy 2078	gaggcacactagatagaacattatgtctatgaagaaagtgtacagcccttaagca 2137
Dd 9	GAGGACACACTAGAGTAGAAACTTTATGTGTATGGGAATGTGTACGCCCTTATGCCA 68
Oy 2138	ggagtcatactcatcatagaccacagagacacacacagtcgttgcttttccaagcatt 2197
Dd 69	GGAGTCATCTCTCATCAGACACACCAGAGCACACACAGCGTGCGCTTTTCAGGCATT 128
Oy 2198	gctaataaccaagtgaggacaacctctgttgtgatatacgatgagagctgtactgttaaga 2257
Dd 129	GCTATATCCAAAGAAGGAGAACATTTCTGTCTGTAATTATGCATGAGACTGTACTGTAAACA 188
Oy 2258	cctgtatcccatccacccgcgaagagaattgtctgtctatttcaagagccctgccttc 2317
Dd 189	CTTGATATCTCCATCTCACTGACCTGAGAGGAATAATGCTGGCTCATTTTCAAGAGCCCTGGCCTTC 248
Oy 2318	ctcaactgtgaaatgtgtgttctgtggaaccocggtlcaagtaatatagttgycaggaagcagt 2377
Dd 249	CTCACTGTGGATGHTGGGTTGTGGAAACCCGGTCAAGTGATATGATAGTGCCAGAGGAGCACT 308
Oy 2378	caaatgccacagcgagataaggggtgtgttacctgtgtgtaaaccacaaccttaagctgaagaca 2437
Dd 309	CAAAATGCCACAGCCAGCATATAGGGGTGGGTACTGTGTGAACCCCAACCTTAAAGCTGAAGACA 368
Oy 2438	gtcccggcgctaaatcctcatctaacgaaattggaacctgctctccatttgtgtgtctct 2497
Dd 369	GTCGGCGCTAAATCTCTCATACTGAAATTGGAACCTGTCTTCCCAATTTGGTGTGCTTCT 428
Oy 2498	ccgaattgatcccaacccttcaacctatltaagtatacctgtcccttcccaatttglttt 2557
Dd 429	CCGATTTGATCCCAACCCCTTCACACCTATTATTAGCATATACCGCCCTTCCAAATTTGGTTTT 488
Oy 2558	aacatgctgtgtcccaaccttttttagtgtgtgctcttgcatatactaaataatgaatcaagctgt 2617
Dd 489	ACACGTGCTGTGCCACCTTTTGTAGTGGTGCCCTTTTGCACTATCAAAATCAGTCAACAGTGT 548
Oy 2618	attccccattcttgagcccataaagaccca 2648
Dd 549	ATTCCCCTATTTCTGAGCCCATATAAAGACCCA 579
RESULT 9	
LOCUS AC024911	130351 bp DNA linear HTG 02-MAR-2000
DEFINITION	Homo sapiens chromosome 3p clone RP11-420a23, WORKING DRAFT

ACCESSION	SEQUENCE, 32 unordered pieces.
VERSION	AC024911.1 GI:7143406
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 130351) Lu,G., Zhao,Y., Lin,W., Dong,H., Wan,M., Xu,S., Gu,W., Tu,Y., Jia,J., Wu,C., Zhang,C., Zhong,M., Zhou,Y., Ren,S., Fu,G., Chen,Z and Huang,M.
TITLE	Direct Submission
JOURNAL	Submitted (02-MAR-2000) Genomic Dept., Chinese National Human Genome Center at Shanghai, 315 Guo Shou Jing Road, Shanghai, Shanghai 201203, P. R. China
COMMENT	-----Genome Center Informations----- Center: Chinese National Human Genome Center at Shanghai Center Code: CHGC Web site: http://www.chgc.sh.cn Email: mhuang@chgc.sh.cn or fuyang@chgc.sh.cn -----End Genome Center Informations----- * NOTE: This is a 'working draft' sequence. It currently * consists of 32 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * * 1 2267: contig of 2267 bp in length * * 2268 gap of unknown length * 4378: contig of 2111 bp in length * * gap of unknown length * 4379 6650: contig of 2272 bp in length * * gap of unknown length * 6651 8791: contig of 2141 bp in length * * gap of unknown length * 8792 11482: contig of 2691 bp in length * * gap of unknown length * 11483 13872: contig of 2390 bp in length * * gap of unknown length * 13873 16745: contig of 2873 bp in length * * gap of unknown length * 16746 18965: contig of 2220 bp in length * * gap of unknown length * 18966 21604: contig of 2639 bp in length * * gap of unknown length * 21605 24066: contig of 2462 bp in length * * gap of unknown length * 24067 26444: contig of 2378 bp in length * * gap of unknown length * 26445 29243: contig of 2799 bp in length * * gap of unknown length * 29244 31995: contig of 2752 bp in length * * gap of unknown length * 31996 34538: contig of 2543 bp in length * * gap of unknown length * 34539 37970: contig of 3432 bp in length * * gap of unknown length * 37971 41593: contig of 3623 bp in length * * gap of unknown length * 41594 44802: contig of 3209 bp in length * * gap of unknown length * 44803 48691: contig of 3889 bp in length * * gap of unknown length * 48692 52259: contig of 3568 bp in length * * gap of unknown length * 52260 55732: contig of 3473 bp in length * * gap of unknown length * 55733 57756: contig of 2024 bp in length * * gap of unknown length * 57757 61389: contig of 3633 bp in length

```
* * *
* 61390 65398: contig of 4009 bp in length
* * *
* 65399 69740: contig of 4342 bp in length
* * *
* 69741 74577: gap of unknown length
* * *
* 74578 80165: gap of unknown length
* * *
* 80166 86055: gap of unknown length
* * *
* 86056 89215: gap of unknown length
* * *
* 89216 94479: gap of unknown length
* * *
* 94480 103136: gap of unknown length
* * *
* 103137 113254: contig of 10118 bp in length
* * *
* 113255 130351: contig of 17097 bp in length.
Location/Qualifiers
  source
    1.130351
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="3p"
    /clone="RP11-420a23"
BASE COUNT 38785 a 26050 c 26300 g 39176 t 40 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 attgagaacctgtctccattgtgtgtctctc 2498
      |||
Db 69153 ATTGAGAACCTGTCTCCATTGTGTCTCTC 69189

RESULT 10
AC096898 161625 bp DNA linear HTG 12-JAN-2002
LOCUS Homo sapiens chromosome 4 clone RP11-618K19, WORKING DRAFT
DEFINITION
AC096898
SEQUENCE 3 unordered pieces.
AC096898.5 GI:18139542
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 161625)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 161625)
TITLE Direct Submission
AUTHORS Waterston,R.H.
JOURNAL Submitted (01-OCT-2001) Genome Sequencing Center, Washington
REFERENCE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jan 12, 2002 this sequence version replaced g1:17921273.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: H_NH0618K19
----- Summary Statistics -----
Sequencing vector: M13; 0%
```

```
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 160310 bases at least Q40
Consensus quality: 160620 bases at least Q30
Consensus quality: 160738 bases at least Q20
Insert size: 164000; agarose-gel
Insert size: 161425; sum-of-contigs
Quality coverage: 7.76 in Q20 bases; sum-of-contigs
Quality coverage: 8.11 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
* 3023: contig of 3023 bp in length
* 3024 3123: gap of unknown length
* 3124 115853: contig of 112730 bp in length
* 115854 115953: gap of unknown length
* 115954 161625: contig of 45672 bp in length.
Location/Qualifiers
  source
    1.161625
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="4"
    /clone="RP11-618K19"
    1.3023
    /note="assembly_name:Contig24
    clone_end:SP6
    vector_side:left"
    3124.115853
    /note="assembly_name:Contig26
    clone_end:T7
    vector_side:right"
    115954.161625
    /note="assembly_name:Contig25"
BASE COUNT 47603 a 31872 c 31765 g 50185 t 200 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 attgagaacctgtctccattgtgtgtctctc 2498
      |||
Db 110301 ATTGAGAACCTGTCTCCATTGTGTCTCTC 110337

RESULT 11
AC034197 164264 bp DNA linear HTG 29-MAY-2000
LOCUS Homo sapiens chromosome 3 clone RP11-72801 map 3p, WORKING DRAFT
DEFINITION
AC034197
SEQUENCE 14 unordered pieces.
AC034197.3 GI:8101273
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 164264)
AUTHORS Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,D., Li,L.,
Li,S., Li,T., Liu,Y., Liu,Y., Liu,X., Liu,Y., Li,W., Li,Y.,
Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
Tao,R., Wang,H., Wang,J., Wang,D., Wang,L., Wang,L., Wang,R.,
```

TITLE
JOURNAL
REFERENCE
AUTHORS

Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y.,
Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H.,
Zhang, L., Zhang, M., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Y.,
Zhang, Z., Zhu, B., Yu, J. and Yang, H.
Chromosome 3p genomic sequence
Unpublished
2 (bases 1 to 164264)

Zeng, Y., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B.,
Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y.,
Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y.,
Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D.,
Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J.
and Yang, H.

Direct Submission
Submitted (05-APR-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P. R. China
On May 29, 2000 this sequence version replaced gi:7644466.

-----Genome Center
Center: Beijing Center
Center code: Beijing
Website: http://hgc.igtp.ac.cn
http://www.genomics.org.cn
Contact: hgc@igtp.ac.cn
-----Project Information
Center project name: 1% project
Center clone name: RP11-72801

-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; Et 5% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 159801 bases at least Q40
Consensus quality: 163151 bases at least Q30
Consensus quality: 165125 bases at least Q20
Insert size: 158755; sum-of-contigs
Quality coverage: 4.57x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2043: contig of 2043 bp in length
* 2044 2143: gap of unknown length
* 2144 4407: contig of 2264 bp in length
* 4408 4507: gap of unknown length
* 4508 5590: contig of 1083 bp in length
* 5591 5690: gap of unknown length
* 5691 8894: contig of 3204 bp in length
* 8895 12938: contig of 3944 bp in length
* 12939 13038: gap of unknown length
* 13039 19151: contig of 6113 bp in length
* 19152 19251: gap of unknown length
* 19252 24784: contig of 5533 bp in length
* 24785 24885: gap of unknown length
* 24886 32377: contig of 7493 bp in length
* 32378 32478: gap of unknown length
* 32479 41545: contig of 9068 bp in length
* 41546 41645: gap of unknown length
* 41646 50878: contig of 9233 bp in length
* 50879 50979: gap of unknown length
* 50980 62812: contig of 11834 bp in length
* 62813 62912: gap of unknown length
* 62913 88788: contig of 25876 bp in length
* 88789 88888: gap of unknown length
* 88889 115986: contig of 27098 bp in length
* 115987 116086: gap of unknown length
* 116087 164264: contig of 48178 bp in length.

TITLE
JOURNAL
COMMENT

FEATURES
source

Location/Qualifiers
1..164264
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3p"
/map="3p"
/clone="RP11-72801"
1..2043
/note="assembly_name:Contig4"
2144..4407
/note="assembly_name:Contig5"
4508..5590
/note="assembly_name:Contig6
clone_end:T7
vector_side:right"
5591..8894
/note="assembly_name:Contig7"
8895..12938
/note="assembly_name:Contig8"
13039..19151
/note="assembly_name:Contig9"
19252..24784
/note="assembly_name:Contig10"
24885..32377
/note="assembly_name:Contig11"
32478..41545
/note="assembly_name:Contig12"
41646..50878
/note="assembly_name:Contig13"
50979..62812
/note="assembly_name:Contig14"
62913..88788
/note="assembly_name:Contig15"
88889..115986
/note="assembly_name:Contig16"
116087..164264
/note="assembly_name:Contig17"
BASE COUNT 46659 a 32302 c 32891 g 51080 t 1332 others
ORIGIN

Query Match 1.3% Score 37; DB 2; Length 164264;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 attggaaccgtctccattgtgtgcttcctc 2498
|||||
Db 55166 ATTGAGACCTGCTCTCCATTGTGTGCTTCCTC 55202

RESULT 12
AC069265
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-208K18 map 3p, WORKING DRAFT
SEQUENCE 7 unordered pieces.
ACCESSION AC069265
VERSION AC069265.2 GI:8101151
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 165379)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Dong, W., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H.,
Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, G., Li, J., Li, L.,
Li, S., Li, T., Liu, Y., Liu, N., Song, S., Sun, M., Sun, W., Sun, Y.,
Tao, R., Wang, H., Wang, J., Wang, L., Wang, L., Wang, R.,
Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y.,
Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H.,
Zhang, L., Zhang, M., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Y.,
Zhang, Z., Zhu, B., Yu, J. and Yang, H.

TITLE	Chromosome 3p genomic sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 165379)
AUTHORS	Wang,L., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D., Huang,F., Zhang,G., Li,J., Bian,X., Zhang,W., Li,L., Feng,X., Yu,J. and Yang,H.
TITLE	Direct Submission
JOURNAL	Submitted (23-MAY-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, 100101, P.R.China
COMMENT	On May 29, 2000 this sequence version replaced gi:8039668. -----Genome Center Center:Beijing Center Center code:Beijing Website:http://hg.c.igtp.ac.cn http://www.genomics.org.cn Contact:hgcelgtp.ac.cn -----Project Information Center project name:1% project Center clone name: RP11-208K18 -----Summary Statistics Sequencing vector: pUC18; 100% of reads Chemistry: Dye-terminator; ET 55% of reads Chemistry: Dye-terminator Big Dye; 45% of reads Assembly program: Phrap; version 0.990329 Consensus quality: 164602 bases at least Q40 Consensus quality: 165885 bases at least Q20 Consensus quality: 166968 bases at least Q20 Insert size: 162389; sum-of-contigs Quality coverage: 6.82x in Q20 bases;sum-of-contigs ----- * NOTE: This is a 'working draft' sequence. It currently * consists of 7 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * * 1 1874: contig of 1874 bp in length * 1875 1974: gap of unknown length * 1975 8441: contig of 6467 bp in length * 8442 8541: gap of unknown length * 8542 20983: contig of 12442 bp in length * 20984 21083: gap of unknown length * 21084 38290: contig of 17207 bp in length * 38291 38390: gap of unknown length * 38391 56760: contig of 18370 bp in length * 56761 56860: gap of unknown length * 56861 93526: contig of 36666 bp in length * 93527 93626: gap of unknown length * 93627 165379: contig of 71753 bp in length. Location/Qualifiers source 1. 165379 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="3" /map="3p" /clone="RP11-208K18" 1. 1874 /note="assembly_name:Contig3" 1975. 8441 /note="assembly_name:Contig4" 8542. 20983 /note="assembly_name:Contig5" 21084. 38290 /note="assembly_name:Contig6" 38391. 56760 misc_feature 56861. 93526 misc_feature

misc_feature		/note="assembly_name:Contig8"	
93627..165379		/note="assembly_name:Contig9"	
BASE COUNT	5145 a 32065 c 32897 g 48394 t	608 others	
ORIGIN			
Query Match	1.3%	Score 37;	DB 2; Length 165379;
Best Local Similarity	100.0%	Pred. No. 2,4e-09;	
Matches 37; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy 2462	attgagaccgtctccattggtgcttcctc 2498		
Db 83850	ATTGAGAACCTGCTTCCATTGGTGCTTCTC 83886		
RESULT 13			
LOCUS	AC026219	184375 bp	DNA linear HTG 29-MAY-2000
DEFINITION	Homo sapiens chromosome 3 clone RP11-815J19 map 3p, WORKING DRAFT		
ACCESSION	AC026219	SEQUENCE, 56 unordered pieces.	
VERSION	AC026219.2	GI:8101287	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 184375)		
	Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y., Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Yu,T. and Yang,H.		
TITLE	Chromosome 3p genomic sequence		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 184375)		
AUTHORS	Bao,N., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Kang,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D., Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J. and Yang,H.		
	Direct Submission		
	Submitted (21-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P R China		
COMMENT	On May 29, 2000 this sequence version replaced gi:7272036.		
	-----Genome Center		
	Center:Beijing Center		
	Center code:Beijing		
	Website:http://hgsc.igtp.ac.cn		
	http://www.genomics.org.cn		
	Contact:hgsc@igtp.ac.cn		
	----- Project Information		
	Center project name:1% project		
	Center clone name: RP11-815J19		
	----- Summary Statistics		
	Sequencing vector: pUC18: 100% of reads		
	Chemistry: Dye-terminator: ET 55% of reads		
	Chemistry: Dye-terminator: Big Dye: 45% of reads		
	Assembly program: Phrap: version 0.990329		
	Consensus quality: 146985 bases at least Q40		
	Consensus quality: 167323 bases at least Q30		
	Consensus quality: 182494 bases at least Q20		
	Insert size: 122192; sum-of-contigs		
	Quality coverage: 3.83x in Q20 bases;sum-of-contigs		

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 56 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      1389: contig of 1389 bp in length
1390      1489: gap of unknown length
1490      3602: contig of 2113 bp in length
3603      3702: gap of unknown length
3703      6147: contig of 2445 bp in length
6148      6247: gap of unknown length
6248      7841: contig of 1594 bp in length
7842      7941: gap of unknown length
7942      9569: contig of 1628 bp in length
9570      9669: gap of unknown length
9670      11970: contig of 2301 bp in length
11971      12070: gap of unknown length
12071      13209: contig of 1139 bp in length
1210      13309: gap of unknown length
13310      14906: contig of 1597 bp in length
14907      15006: gap of unknown length
15007      16915: contig of 1909 bp in length
16916      17015: gap of unknown length
17016      18655: contig of 1640 bp in length
18656      18755: gap of unknown length
18756      20974: contig of 2219 bp in length
20975      21074: gap of unknown length
21075      22409: contig of 1335 bp in length
22410      22509: gap of unknown length
22510      24337: contig of 1828 bp in length
24338      24437: gap of unknown length
24438      25826: contig of 1389 bp in length
25827      25926: gap of unknown length
25927      27769: contig of 1843 bp in length
27770      27869: gap of unknown length
27870      29441: contig of 1572 bp in length
29442      29541: gap of unknown length
29542      32494: contig of 2953 bp in length
32495      32594: gap of unknown length
32595      34633: contig of 2039 bp in length
34634      34733: gap of unknown length
34734      37159: contig of 2426 bp in length
37160      37259: gap of unknown length
37260      39230: contig of 1971 bp in length
39231      39330: gap of unknown length
39332      40642: contig of 1312 bp in length
40643      40742: gap of unknown length
40743      42739: contig of 1997 bp in length
42740      42839: gap of unknown length
42840      45245: contig of 2406 bp in length
45246      45345: gap of unknown length
45346      47083: contig of 1738 bp in length
47084      47183: gap of unknown length
47184      49945: contig of 2762 bp in length
49946      50045: gap of unknown length
50046      51544: contig of 1499 bp in length
51545      51644: gap of unknown length
51645      53891: contig of 2247 bp in length
53892      53991: gap of unknown length
53992      57221: contig of 3320 bp in length
57222      57321: gap of unknown length
57322      60016: contig of 2695 bp in length
60017      60116: gap of unknown length
60117      62563: contig of 2447 bp in length
62564      62663: gap of unknown length
62664      64926: contig of 2263 bp in length
64927      65026: gap of unknown length
65027      68373: contig of 3347 bp in length
68374      68473: gap of unknown length
68474      71700: contig of 3227 bp in length

```

```

71701      71800: gap of unknown length
71801      74355: contig of 2555 bp in length
74356      74455: gap of unknown length
74456      77266: contig of 2811 bp in length
77267      77366: gap of unknown length
77367      80106: contig of 2740 bp in length
80107      80206: gap of unknown length
80207      82763: contig of 2557 bp in length
82764      82863: gap of unknown length
82864      85978: contig of 3015 bp in length
85979      86748: gap of unknown length
86749      88748: contig of 2770 bp in length
88749      93166: contig of 4318 bp in length
93167      93266: gap of unknown length
93267      96634: contig of 3368 bp in length
96635      96734: gap of unknown length
96735      100449: contig of 3715 bp in length
100450      100549: gap of unknown length
100550      103083: contig of 2534 bp in length
103084      103183: gap of unknown length
103184      107029: contig of 3846 bp in length
107030      107129: gap of unknown length
107130      110324: contig of 3195 bp in length
110325      110424: gap of unknown length
110425      114667: contig of 4043 bp in length
114668      114567: gap of unknown length
114568      119345: contig of 4778 bp in length
119346      119445: gap of unknown length
119446      122439: contig of 2994 bp in length
122440      122539: gap of unknown length
122540      127598: contig of 5059 bp in length
127599      127698: gap of unknown length
127699      133792: contig of 6094 bp in length
133793      133892: gap of unknown length
133893      140225: contig of 6333 bp in length
140226      140325: gap of unknown length
140326      144336: contig of 4011 bp in length
144337      144436: gap of unknown length
144437      151171: contig of 6735 bp in length
151172      151271: gap of unknown length
151272      159310: contig of 8039 bp in length
159311      159410: gap of unknown length
159411      168312: contig of 8902 bp in length
168313      168412: gap of unknown length
168413      184375: contig of 15963 bp in length.

```

FEATURES

SOURCE

```

1. 184375
   location/Qualifiers
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="3"
   /map="3p"
   /clone="Rp11-815J19"
1. 1389
   /note="assembly_name:Contig35"
1490. 3602
   /note="assembly_name:Contig37"
3703. 6147
   /note="assembly_name:Contig42"
6248. 7841
   /note="assembly_name:Contig45"
7942. 9569
   /note="assembly_name:Contig48"
9670. 11970
   /note="assembly_name:Contig50"
12071. 13209
   /note="assembly_name:Contig51"
13310. 14906
   /note="assembly_name:Contig52"
15007. 16915
   /note="assembly_name:Contig53"
17016. 18655
   misc_feature

```

Query Match 1.3%; Score 37; DB 2; Length 184375;
 Best Local Similarity 100.0%; Pred. 2.3e-09;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 attgagaacctgtctccattgtgtgtcttcctc 2498
 |||||||
 Db 158390 ATTGAGAACCTGCTCCATTGTGTGTCTTCCTC 158426

RESULT 14
 AC027579 148270 bp DNA linear HTG 23-JUN-2000
 LOCUS Homo sapiens chromosome 17 clone RP11-177D1 map 17, WORKING DRAFT
 DEFINITION
 AC027579 13 unordered pieces.
 AC027579.2 GI:8671976
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 148270)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 17, clone RP11-177D1
 Unpublished
 2 (bases 1 to 148270)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
 Boguslavskiy,L., Bouknight,B., Brown,A., Burkett,G.,
 Campolavio,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Coyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., Laroque,K., Lamazares,R., Lander,S., Lechoczky,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
 Meldrum,J., Meneau,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testafaye,S., Theodore,J., Tirrell,A., Travers,M., Triggilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,W.
 Direct Submission
 Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 23, 2000 this sequence version replaced gi:7342324.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L9001
 Center clone name: 177_D_1

----- Summary Statistics
 Sequencing vector: M13; M77815: 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 141141 bases at least Q40
 Consensus quality: 144792 bases at least Q30
 Consensus quality: 146242 bases at least Q20
 Insert size: 151000; agarose-fp
 Insert size: 147070; sum-of-ctrls
 Quality coverage: 4.7 in Q20 bases; agarose-fp
 Quality coverage: 4.8 in Q20 bases; sum-of-ctrls

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 13 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of 'N', but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 131: contig of 131 bp in length
 132 231: gap of 100 bp
 232 4256: contig of 4025 bp in length
 4257 4356: gap of 100 bp
 4357 6715: contig of 2359 bp in length
 6716 6815: gap of 100 bp
 6816 12854: contig of 6039 bp in length
 12855 12954: gap of 100 bp
 12955 18900: contig of 5946 bp in length
 18901 19000: gap of 100 bp
 19001 27560: contig of 8560 bp in length
 27561 27660: gap of 100 bp
 27661 41935: contig of 14275 bp in length
 41936 42035: gap of 100 bp
 42036 58175: contig of 16140 bp in length
 58176 58275: gap of 100 bp
 58276 70930: contig of 12655 bp in length
 70931 71030: gap of 100 bp
 71031 87632: contig of 16602 bp in length
 87633 87732: gap of 100 bp
 87733 104759: contig of 17027 bp in length
 104760 104859: gap of 100 bp
 104860 123863: contig of 19004 bp in length
 123864 123963: gap of 100 bp
 123964 148270: contig of 24307 bp in length.

Location/Qualifiers
 1. 148270
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-177D1"
 /clone_1lb="RP11 Human Male BAC"
 1. 131
 /note="assembly-fragment
 clone_end:T7
 vector_side:right"
 232. 4256
 /note="assembly-fragment"
 4357. 6715
 /note="assembly-fragment
 clone_end:SP6
 vector_side:right"
 6816. 12854
 /note="assembly-fragment"
 12955. 18900
 /note="assembly-fragment"
 19001. 27560
 /note="assembly-fragment"
 27661. 41935
 /note="assembly-fragment"
 42036. 58175
 /note="assembly-fragment"
 58276. 70930
 /note="assembly-fragment"
 71031. 87632
 /note="assembly-fragment"
 87733. 104759
 /note="assembly-fragment"
 104860. 123863
 /note="assembly-fragment"
 123964. 148270
 /note="assembly-fragment"

BASE COUNT 40333 a 32099 c 32432 g 42206 t 1200 others

ORIGIN

Query Match 1.3%; Score 36; DB 2; Length 148270;
 Best Local Similarity 100.0%; Pred. No. 9,1e-09;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 atgagaaacctgtccattgtgtgtcttct 2497
 ||||||||||||||||||||||||||||||||||

DB 76167 ATTGAGAACCTGTCTCCATTGTGTCTTCC 76202

RESULT 15

AC022727

LOCUS AC022727 174405 bp DNA linear HTG 26-MAY-2000
 DEFINITION Homo sapiens chromosome 18 clone RP11-308J14 map 18, WORKING DRAFT
 SEQUENCE, 11 unordered pieces.

AC022727

AC022727.4 GI:8072614

HTG: HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

SOURCE

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

1 (bases 1 to 174405)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 18, clone RP11-308J14
 Unpublished
 2 (bases 1 to 174405)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckert,R., Beda,F.,
 Boguslavsky,L., Bouckhgalter,B., Brown,A., Burkett,G., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dearlano,K., Dewar,K., Domino,M., Doyle,M., Feneslor,J.,
 Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
 MacDonald,P., Marquis,N., McEwan,P., McGuirk,A., McKernan,K.,
 Mcneeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
 Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
 Tittel,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
 Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 25, 2000 this sequence version replaced g1:7658384.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L5488

Center clone name: 308_J14

Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 169089 bases at least Q40

Consensus quality: 172164 bases at least Q40

Insert size: 177000; agarose-1p

Quality coverage: 4.8 in Q20 bases; agarose-1p

Quality coverage: 4.9 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 6450: contig of 6450 bp in length
 * 6451 6550: gap of 100 bp
 * 6551 12911: contig of 6361 bp in length
 * 12912 13011: gap of 100 bp
 * 13012 22057: contig of 9046 bp in length
 * 22058 22157: gap of 100 bp
 * 22158 32841: contig of 10684 bp in length
 * 32842 32941: gap of 100 bp
 * 32942 48367: contig of 15426 bp in length
 * 48368 48467: gap of 100 bp
 * 48468 63919: contig of 15452 bp in length
 * 63920 64019: gap of 100 bp
 * 64020 78259: contig of 14240 bp in length
 * 78260 78359: gap of 100 bp
 * 78360 98415: contig of 20056 bp in length
 * 98416 98515: gap of 100 bp
 * 98516 121876: contig of 23361 bp in length
 * 121877 121976: gap of 100 bp
 * 121977 144489: contig of 22513 bp in length
 * 144490 144589: gap of 100 bp
 * 144590 174405: contig of 29816 bp in length.

FEATURES

source

1. 174405

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="18"

/map="18"

/clone="RP11-308J14"

/clone_id="RP11-308J14 Human Male BAC"

1. 6450

/note="assembly-fragment"

6551. 12911

/note="assembly-fragment"

13012. 22057

/note="assembly-fragment"

22158. 32841

/note="assembly-fragment"

32942. 48367

/note="assembly-fragment"

48468. 63919

/note="assembly-fragment"

64020. 78259

/note="assembly-fragment"

vector_end:SP6

vector_side:right"

78360. 98415

/note="assembly-fragment"

clone_end:T7

vector_side:left"

98516. 121876

/note="assembly-fragment"

121977. 144489

/note="assembly-fragment"

144590. 174405

/note="assembly-fragment"

BASE COUNT 46895 a 38223 c 37326 g 50956 t 1003 others.

ORIGIN

Query Match 1.3%; Score 36; DB 2; Length 174405;
 Best Local Similarity 100.0%; Pred. No. 9e-09;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ||||||||||||||||||||||||||||||||||

Sat May 18 14:46:05 2002

us-09-898-556a-3.rge

Page 19

Db 127497 ATGAGAACCTGCTTCCATTGTTGGTGGCTTCCCT 127532

Search completed: May 17, 2002, 11:55:34
Job time: 75409 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 14:43:43 ; Search time 539.13 Seconds
(without alignments)
8827.717 Million cell updates/sec

Title: US-09-898-556a-3
Perfect score: 2772
Sequence: 1 cagcgcgcttaagctggtg.....ttctaccatcttcacct 2772

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N.Geneseq_032802.*

```
1: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1577	56.9	2614	23	AA574823	DNA encoding novel
2	571	20.6	678	22	AA574823	Human breast cancer
3	525	18.9	653	23	AA568256	DNA encoding novel
4	231	8.3	983	22	AA56286	Human cDNA encoding
5	99	3.6	394	22	AA183880	Human polynucleotide
6	69	2.5	69	22	AA513031	DNA encoding zinc
7	69	2.5	69	22	AA513088	DNA encoding zinc
8	49	1.8	821	23	AA568257	DNA encoding novel
9	29	1.0	800	22	AA53832	Human colon cancer

10	29	1.0	2349	23	AA590429	DNA encoding novel
11	29	1.0	3582	22	AA158664	Human polynucleotide
12	29	1.0	3582	22	AA158664	Human polynucleotide
13	28	1.0	159	22	AA100216	Human reproductive
14	27	1.0	1183	22	AA525949	Human cDNA encoding
15	27	1.0	2717	22	AA526635	Human genomic DNA
16	27	1.0	3099	20	AA64361	Human stem cell z1
17	25	0.9	51	23	AB100474	Human silent nonco
18	25	0.9	51	23	AB100878	Human amino acid c
19	25	0.9	753	22	AAK92290	Human cDNA 5'-end
20	25	0.9	763	22	AAK93811	Human cDNA clone r
21	25	0.9	814	22	AA195339	Human neuroblastom
22	25	0.9	871	22	ABA08823	Human PRO1847 homo
23	23	0.9	2200	22	AAK94575	Human full-length
24	23	0.8	688	22	AAK91560	Human cDNA 5'-end
25	23	0.8	688	22	AAK93129	Human cDNA clone r
26	23	0.8	1512	22	ABA50438	Human breast cell
27	23	0.8	1512	22	ABA68388	Human foetal liver
28	23	0.8	1512	22	ABA35382	Probe #13848 for g
29	23	0.8	1512	22	AAK16761	Human brain expres
30	23	0.8	1512	22	AAK42537	Human bone marrow
31	23	0.8	1512	22	AA123284	Probe #13217 for g
32	23	0.8	1512	22	AA148604	Probe #17290 used t
33	23	0.8	1512	22	AA108925	Probe #8916 used t
34	23	0.8	1623	22	AAK94635	Human full-length
35	23	0.8	1965	22	ABA45305	Human breast cell
36	23	0.8	1965	22	ABA55794	Human foetal liver
37	23	0.8	1965	22	ABA25474	Probe #3940 for ge
38	23	0.8	1965	22	AAK04015	Human brain expres
39	23	0.8	1965	22	AAK29501	Human bone marrow
40	23	0.8	1965	22	AA114072	Probe #4005 for ge
41	23	0.8	1965	22	AA135452	Probe #4138 used t
42	23	0.8	1965	22	AA103925	Probe #3916 used t
43	23	0.8	2634	23	AA581041	DNA encoding novel
44	22	0.8	439	22	AAK67397	Novel human polyn
45	22	0.8	484	21	AAK79234	Human lung tumour

ALIGNMENTS

RESULT 1
ID AAS74823 standard; cDNA; 2614 BP.
XX AAS74823:
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #10627.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
(HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB: ABG10636.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

Claim 1: SEQ ID No 10627; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX - Sequence 2614 BP; 745 A; 611 C; 660 G; 598 T; 0 other:

Query Match 56.9%; Score 1577; DB 23; Length 2614;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1817; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 427 cagaatcgaagcaaaattcaactagtcctcctgcctctgtattttctcagtcacg 486
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 229 cagaatcgaagcaaaattcaactagtcctcctgcctctgtattttctcagtcacg 288
OY 487 aagctctcagccaacatgtgtgctgagtcacatctcctcagctgtttccaagttagg 546
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 289 aagctctcagccaacatgtgtgctgagtcacatctcctcagctgtttccaagttagg 348
OY 547 cagg-aaatcctctcactctggaaacacatctcagaagatcagaacaaacagcagat 605
DB ||||| ||||||||||||||||||||||||||||||||||||||||||||
DB 349 caggaaatcctctcactctggaaacacatctcagaagatcagaacaaacagcagat 408
OY 606 ccatctctcttaagtggaagaacagatgattcaagagggaagactccagactctg 665
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 409 ccatctctcttaagtggaagaacagatgattcaagagggaagactccagactctg 468
OY 666 ttggagagtagtaagcaaaatgtgcaacttcaagagcacttccagcccaactgaagaaca 725
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 469 ttggagagtagtaagcaaaatgtgcaacttcaagagcacttccagcccaactgaagaaca 528
OY 726 cagcagcagcagtcgaaggaagaacaacagtggtgatatagggtccagccctgaacgg 785
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 529 cagcagcagcagtcgaaggaagaacaacagtggtgatatagggtccagccctgaacgg 588
OY 786 agggcagagtcagaggaagaacagacatgattgattgattgaaggtctcagatttga 845
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 589 agggcagagtcagaggaagaacagacatgattgattgattgaaggtctcagatttga 648
OY 846 gaatacaaatatgaagagcttggccaagcttatacaagagtcacaactccttgctc 905
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 649 gaatacaaatatgaagagcttggccaagcttatacaagagtcacaactccttgctc 708
OY 906 cagaagacacaaactgggagagacacctatctactgagtgggagacagcttggc 965
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 709 cagaagacacaaactgggagagacacctatctactgagtgggagacagcttggc 768
OY 966 agtatgtcagtcctatcaaaaaaaccaagagacacactctgggggaagccttatgtg 1025
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 769 agtatgtcagtcctatcaaaaaaaccaagagacacactctgggggaagccttatgtg 828
OY 1026 aggaatctgggcggaagcttaccgtggaagtcaaaactcgatcacacatcagaagagcacac 1085
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 829 aggaatctgggcggaagcttaccgtggaagtcaaaactcgatcacacatcagaagagcacac 888
OY 1086 tcaagggaagaacacttatgtgtcagaagattgtgacagagcgttacttggaaagtcagac 1145
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 889 tcaagggaagaacacttatgtgtcagaagattgtgacagagcgttacttggaaagtcagac 948
OY 1146 ctcttaccatcagcagcagac 1205
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
OY 1206 cagaagctttagcctgaagtcacaactcttaccacacagagggcgacacacacacacacac 1265
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1009 cagaagctttagcctgaagtcacaactcttaccacacagagggcgacacacacacacacac 1068
OY 1266 ccttatgttcagaaggaatgtgtgctgttcgcacagcattcacacctgttcagacac 1325
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1069 ccttatgttcagaaggaatgtgtgctgttcgcacagcattcacacctgttcagacac 1128
OY 1326 aagagacacatctcagggagagagccttacatttgcagggagtgatgacagagctttagc 1385
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1129 aagagacacatctcagggagagagccttacatttgcagggagtgatgacagagctttagc 1188
OY 1386 cagaagtcacacactatcagacacacttaagagacacacacacacacacacacacacac 1445
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1189 cagaagtcacacactatcagacacacttaagagacacacacacacacacacacacacac 1248
OY 1446 acagaatctgggcgtcactcttagctggaatcaaaactcacaacacacacacacacacacac 1505
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1249 acagaatctgggcgtcactcttagctggaatcaaaactcacaacacacacacacacacac 1308
OY 1506 tcaagggttaaaccttatgctgcctgagtgagtgaggcgctgtttagcctgaagtcacac 1565
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1309 tcaagggttaaaccttatgctgcctgagtgagtgaggcgctgtttagcctgaagtcacac 1368
OY 1566 cttaacaacaac 1624
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1369 cttaacaacaac 1428
OY 1625 gcgagagcttaccgcggaatacaacccctgaatcaagcacaagagacacacacacacacac 1684
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1429 gcgagagcttaccgcggaatacaacccctgaatcaagcacaagagacacacacacacac 1488
OY 1685 gccatttgatgtgtgctgagtggtggaagagccttataagatgacacacacacacacac 1744
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1489 gccatttgatgtgtgctgagtggtggaagagccttataagatgacacacacacacacac 1548
OY 1745 ccagagagacacatcagggggaagaagccttatagtgcagggagtggtgcagaagagtttcg 1804
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1549 ccagagagacacatcagggggaagaagccttatagtgcagggagtggtgcagaagagtttcg 1608
OY 1805 gcaagaagccttaacctgttttaaggcacagaagggcacacacacacacacacacacacacac 1864
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1609 gcaagaagccttaacctgttttaaggcacagaagggcacacacacacacacacacacacac 1668
OY 1865 gttgtgcaagagccttctgtcctaagtttaacctcttaataaacacagagagacacacacac 1924
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1659 gttgtgcaagagccttctgtcctaagtttaacctcttaataaacacagagagacacacacac 1728
OY 1925 ggggaagcctcatgtgtcagaaggagtggtgcaagagcctttagccggcagtcacacacac 1984
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1729 ggggaagcctcatgtgtcagaaggagtggtgcaagagcctttagccggcagtcacacacac 1788
OY 1985 tagacacagagagacacatcagaagagagagccttataattgcaagaaagtgtgacagggg 2044
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1789 tagacacagagagacacatcagaagagagagccttataattgcaagaaagtgtgacagggg 1848
OY 2045 cttagtctggaagtcacaactctatcagacacatcagaagagacacacacacacacacacacac 2104
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1849 cttagtctggaagtcacaactctatcagacacatcagaagagacacacacacacacacacacac 1908

(II). (I) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pat_sequences](http://wipo.int/pub/published_pat_sequences).

Query Match	18.9%	Score 525;	DB 23;	Length 653;
Best Local Similarity	99.8%	Pred. No. 3.3e-248;		
Matches 575; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

OY	1004	tcgggggaagagccttatgtgtgcaaggaaatgtggcgagagcttaccgttggaagtcacaac	106
Db	52	tcgggggaagagccttatgtgtgcaaggaaatgtggcgagagcttaccgttggaagtcacaac	111
OY	1064	gaccacaatcatagagagacacactcaaggaggaacacttatgtgtgcaagatctgtgagac	112
Db	112	gaccacaatcatagagagacacactcaaggaggaacacttatgtgtgcaagatctgtgagac	171
OY	1124	agccttactcttgaagtcgaacctctttacacatcagcgagacacacttcaggctcaagcc	118
Db	172	agccttactcttgaagtcgaacctctttacacatcagcgagacacacttcaggctcaagcc	231
OY	1184	ttatgtgtgcaaggaaatgtgtggcagaagctttagccctgaagtcacaactctattaccaca	124
Db	232	ttatgtgtgcaaggaaatgtgtggcagaagctttagccctgaagtcacaactctattaccaca	291
OY	1244	gaggcgacacactctggggagagaagccttatgtgttcagagaaatgtgtgcgttcgcga	130
Db	292	gaggcgacacactctggggagagaagccttatgtgttcagagaaatgtgtgcgttcgcga	351
OY	1304	gcattcacacctgtgtcagacacaagaagagacacattcaaggagagaagccttatcttgag	136
Db	352	gcattcacacctgtgtcagacacaagaagagacacattcaaggagagaagccttatcttgag	411
OY	1364	ggagttgtgagaagaagcctttgccaagaatgcacacctcatcagaacacttaagacacac	142
Db	412	ggagttgtgagaagaagcctttgccaagaatgcacacctcatcagaacacttaagacacac	471
OY	1424	aggagagaaagccttatgtatgcacagaatgtgtgggtcactttagctggaatccaacct	148
Db	472	aggagagaaagccttatgtatgcacagaatgtgtgggtcactttagctggaatccaacct	531
OY	1484	caaaacacaccagaagagacacactcaggggttaaacttatgtctgccttgaatctgcygca	154
Db	532	caaaacacaccagaagagacacactcaggggttaaacttatgtctgccttgaatctgcygca	591
OY	1544	gtgcttagcctgaagtcacaacttaacaaaccca	1579
Db	592	gtgcttagcctgaagtcacaacttaacaaaccca	627

RESULT	4
AAAS26286	
ID	AAAS26286 standard; cDNA; 983 BP.
XX	
AC	AAAS26286;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human CDNA encoding a novel secreted protein, Seq ID 465
XX	

KW	Human; immunosuppressive; antiarthritic; ss; antihematic;
KW	neurotrophic; cartilag; vasotropic; cerebroprotective; nootropic;
KW	neuroprotective; antibacterial; vitucide; fungicide; ophthalmological;
KW	vulnary; secreted protein; rheumatoid arthritis;
KW	hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW	cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW	nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW	corneal infection; wound healing; epithelial cell proliferation;
KW	skin ageing; food additive; preservative; antiproliferative.
XX	
OS	Homo sapiens.
XX	
PN	WO200155322-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01341.
XX	
XX	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.
PR	26-JUL-2000; 2000US-0220964.
PR	14-AUG-2000; 2000US-0224518.
PR	14-AUG-2000; 2000US-0224519.
PR	14-AUG-2000; 2000US-0225213.
PR	14-AUG-2000; 2000US-0225214.
PR	14-AUG-2000; 2000US-0225266.
PR	14-AUG-2000; 2000US-0225267.
PR	14-AUG-2000; 2000US-0225268.
PR	14-AUG-2000; 2000US-0225270.
PR	14-AUG-2000; 2000US-0225447.
PR	14-AUG-2000; 2000US-0225757.
PR	14-AUG-2000; 2000US-0225758.
PR	14-AUG-2000; 2000US-0225759.
PR	14-AUG-2000; 2000US-0225759.
PR	18-AUG-2000; 2000US-0226279.
PR	22-AUG-2000; 2000US-0226581.
PR	22-AUG-2000; 2000US-0226688.
PR	22-AUG-2000; 2000US-0227009.
PR	22-AUG-2000; 2000US-0227182.
PR	23-AUG-2000; 2000US-0227009.
PR	30-AUG-2000; 2000US-0228924.
PR	01-SEP-2000; 2000US-0229287.
PR	01-SEP-2000; 2000US-0229343.
PR	01-SEP-2000; 2000US-0229344.
PR	01-SEP-2000; 2000US-0229345.
PR	05-SEP-2000; 2000US-0229509.
PR	05-SEP-2000; 2000US-0229513.
PR	06-SEP-2000; 2000US-0230437.
PR	06-SEP-2000; 2000US-0230438.
PR	06-SEP-2000; 2000US-0231242.
PR	08-SEP-2000; 2000US-0231243.
PR	08-SEP-2000; 2000US-0231244.
PR	08-SEP-2000; 2000US-0231413.
PR	08-SEP-2000; 2000US-0231414.
PR	08-SEP-2000; 2000US-0232080.
PR	08-SEP-2000; 2000US-0232081.
PR	12-SEP-2000; 2000US-0231968.
PR	14-SEP-2000; 2000US-0232397.
PR	14-SEP-2000; 2000US-0232398.
PR	14-SEP-2000; 2000US-0232399.

AA1833890
ID AA183890 standard; cDNA; 394 BP.
AC AA183890;
DT 06-NOV-2001 (first entry)
DE Human polynucleotide SEQ ID NO 3950.
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX MO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-514838/56.
XX P-PSDB; AAO03959.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -
XX
XX Claim 1; SEQ ID NO 3950; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activating
XX activity/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 394 BP; 148 A; 75 C; 93 G; 78 T; 0 other;

	Query Match	3.6%	Score 99;	DB 22;	Length 394;
	Best Local Similarity	99.3%;	Pred. No.	2.6e-38;	
	Matches 149;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	2619	tcccctattctcgagccataaaagaccagactacgtcagttgagagagaatcaccc	2678		
Db	1	ttccccattcttcggccataaagaaccagactacgtcgcgtgtagagagaaaatcac	60		
QY	2679	ctgcgttggagagtctggggacaacctccctgcacatccctctccaactgaagagctgtctcttgy	2738		
Db	61	ctgcgttggagggtctggggacaacctccctgcacatccctctccaactgaagagctgtctcttgy	120		
QY	2739	ctcaataaatctctttctaccacatccctca	2768		
Db	121	ctcaataaatctctttctaccacatccctca	150		

RESULT 6
 ID AAS13031 standard; DNA; 69 BP.
 XX AAS13031.
 AC AAS13031;
 DT 17-DEC-2001 (first entry)
 XX
 XX DNA encoding zinc finger domain TG-ZFD-012.
 XX
 XX zinc finger domain; cancer; human; ds; TG-ZFD-012.
 XX
 OS Homo sapiens.
 XX
 PN W0200160970-A2.
 PD 23-AUG-2001.
 XX
 PF 17-FEB-2001; 2001WO-KR00244.
 XX
 PR 18-FEB-2000; 2000KR-0007730.
 XX
 PA (TOOL-) TOOLGEN INC.
 PI Kim J, Kwon Y, Kim H, Ryu EH, Hwang MS;
 DR WPI; 2001-557644/62.
 DR P-PSDB; AAU08514.
 XX
 XX Identifying a zinc finger domain for e.g. designing new polypeptides
 PT that bind to a specific site on a DNA, comprises expressing hybrid
 PT nucleic acids with a test zinc finger domain in cells -
 PS
 PS Example 22; Page 55; 147pp; English.
 XX
 XX The invention relates to a method of identifying a zinc finger domain
 CC that recognises a target site on a DNA. The method comprises expressing a
 CC hybrid nucleic acids with a test zinc finger domain in cells containing a
 CC reporter construct, where the reporter gene is expressed above a given
 CC level when a transcription factor recognises a recruitment and a target
 CC site of a promoter, and not only the recruitment site of the promoter.
 CC The method is used to: (a) identify a zinc finger domain that recognises
 CC a target site on a DNA; (b) determine whether a test zinc finger domain
 CC recognises a target site on a promoter; (c) generate a nucleic acid that
 CC encodes a chimeric zinc finger protein; and (d) identify DNA sequences
 CC recognised by zinc finger domains. The method can be used to design novel
 CC polypeptides that bind to a specific site on a DNA. The method can
 CC facilitate the customised generation of new polypeptides that can
 CC regulate the expression a selected target e.g. a gene required by a
 CC pathogen can be repressed, a gene required for cancerous growth can be
 CC repressed, or a gene poorly expressed or encoding a mutated protein can
 CC be activated and overexpressed. The method can be used in vivo which
 CC enables identification of polypeptides that bind to a specific site on a
 CC DNA in the intracellular milieu. The present sequence represents the
 CC coding sequence of TG-ZFD-012 zinc finger protein which was used in
 CC the method of the invention.
 XX
 XX Sequence 69 BP; 17 A; 16 C; 20 G; 16 T; 0 other;

Query Match	2.5%	Score 69	DB 22	Length 69
Best Local Similarity	100.0%	Pred. No. 1.ee-23		
Matches 69	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1269	tatgtttgcagagatgtrgggcgtggtctttgcgcagcatctcacactgtctcagacacaag	1328		
Db 1	tatgtttgcagagatgtrgggcgtggtctttgcgcagcatctcacactgtctcagacacaag	60		
QY 1329	agagacacat	1337		
Db 61	agagacacat	69		

AAH3832
ID AAH3832 standard; cDNA; 800 BP.
XX
AC AAH3832;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:888.
XX
KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma, ss.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
PI WPI: 2001-235357/24.
DR P-PSDB; AAG74401.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PS
XX
PS Claim 1: Page 2813-2814; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 800 BP; 240 A; 173 C; 219 G; 166 T; 2 other;

Query Match 1.0%; Score 29; DB 22; Length 800;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcaggaggagaagcc 1687
DB 413 caccagagacacactcaggaggagaagcc 441
|||||

RESULT 10
AAS90429
ID AAS90429 standard; cDNA; 2349 BP.
XX
AC AAS90429;
XX

DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #26233.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
PI WPI: 2001-639362/73.
DR P-PSDB; ABG26242.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1: SEQ ID NO 26233; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2349 BP; 664 A; 547 C; 608 G; 530 T; 0 other;

Query Match 1.0%; Score 29; DB 23; Length 2349;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcaggaggagaagcc 1687
DB 1645 caccagagacacactcaggaggagaagcc 1673
|||||

RESULT 11
AA158664
ID AA158664 standard; cDNA; 3582 BP.
XX
AC AA158664;
XX
DT 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 867.
XX
XX
KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB: AAM39308.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 867; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
SQ Sequence 3582 BP; 1024 A; 826 C; 910 G; 821 T; 1 other;

Query Match 1.0%; Score 29; DB 22; Length 3582;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcaggagagacc 1687
Db 1487 caccagagacacactcaggagagacc 1515

RESULT 12
AA160450/C
ID AA160450 standard; cDNA: 3582 BP.

XX AA160450;
XX
XX
DE 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 4439.
XX
DE Human polynucleotide SEQ ID NO 4439.
XX
KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB: AAM41294.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4439; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
SQ Sequence 3582 BP; 821 A; 908 C; 826 G; 1026 T; 1 other;

Query Match 1.0%; Score 29; DB 22; Length 3582;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcaggagagacc 1687
Db 2096 CACGAGAGACACACTCAGGAGAGACC 2068

RESULT 13
AAL00216
ID AAL00216 standard; cDNA; 159 BP.
XX
AC AAL00216;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen cDNA SEQ ID NO: 217.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN M020015320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0233081.

PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465570/50.
DR P-PSDB; AAM94246.
XX
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
is used in preventing, treating or ameliorating a medical condition -
XX
XX
PS Claim 1; SEQ ID NO 217; 1297bp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a coding sequence of the
CC invention.
XX
SO Sequence 159 BP; 47 A; 33 C; 47 G; 30 T; 2 other:

Query Match 1.0%; Score 28; DB 22; Length 159;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1659 caccagagacacactcagggaggaagc 1686
Db 18 caccagagacacactcagggaggaagc 45

RESULT 14
AAS25949
ID AAS25949 standard; cDNA; 1183 BP.
XX
AC AAS25949;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 128.
XX
XX Human: immunosuppressive; antiarthritic; ss; antirheumatic;
KW cytostatic; cardiant; vasotrophic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischemia; angioneurosis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
XX Homo sapiens.
OS
XX
XX WO200155322-A2.
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001MO-US01341.
PF
XX 31-JAN-2000; 2000US-0179065.
PR
XX 04-FEB-2000; 2000US-0180628.
PR

PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233065.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244618.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251031.
PR 05-DEC-2000; 2000US-0251719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX
DR WPI: 2001-488783/53.
XX P-PSDB; AA015962.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
XX used as food additives or preservatives -

PS Claim 1: SEQ ID No 128; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiodenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention.

Query Match 1.0%; Score 27; DB 22; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2149 tcatcagaccacagagacacacacag 2175
|||||
Db 961 tcatcagaccacagagacacacacag 987

RESULT 15
ID AAS26635 standard; DNA; 2717 BP.
XX
XX AAS26635;
AC
XX
XX 07-NOV-2001 (first entry)
DT
XX
XX Human genomic DNA encoding partial novel secreted protein, Seq ID 1609.
DE
XX
XX Human; immunosuppressive; antiarthritic; ds; antirheumatic;
KW cytoskeletal; cardiant; vasotropic; cerebroprotective; noctropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnerrary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiodenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
XX Homo sapiens.
OS
XX
XX WO200155322-A2.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-0501341.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.

PR	28-JUN-2000	2000US-0214886
PR	30-JUN-2000	2000US-0215133
PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216880
PR	11-JUL-2000	2000US-0217487
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225681
PR	22-AUG-2000	2000US-0226868
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227709
PR	30-AUG-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229282
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	01-SEP-2000	2000US-0229346
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233051
PR	14-SEP-2000	2000US-0233052
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234224
PR	21-SEP-2000	2000US-0234327
PR	21-SEP-2000	2000US-0234328
PR	21-SEP-2000	2000US-0234329
PR	21-SEP-2000	2000US-0234330
PR	25-SEP-2000	2000US-0234988
PR	25-SEP-2000	2000US-0234989
PR	25-SEP-2000	2000US-0235367
PR	25-SEP-2000	2000US-0235368
PR	29-SEP-2000	2000US-0235369
PR	29-SEP-2000	2000US-0235370
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239935
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241781
PR	20-OCT-2000	2000US-0241782
PR	20-OCT-2000	2000US-0241785

CC	20-OCT-2000; 2000US-0241786.
PR	20-OCT-2000; 2000US-0241787.
PR	20-OCT-2000; 2000US-0241808.
PR	20-OCT-2000; 2000US-0241809.
PR	20-OCT-2000; 2000US-0241826.
PR	01-NOV-2000; 2000US-0244617.
PR	08-NOV-2000; 2000US-0246474.
PR	08-NOV-2000; 2000US-0246475.
PR	08-NOV-2000; 2000US-0246476.
PR	08-NOV-2000; 2000US-0246477.
PR	08-NOV-2000; 2000US-0246478.
PR	08-NOV-2000; 2000US-0246523.
PR	08-NOV-2000; 2000US-0246524.
PR	08-NOV-2000; 2000US-0246525.
PR	08-NOV-2000; 2000US-0246526.
PR	08-NOV-2000; 2000US-0246527.
PR	08-NOV-2000; 2000US-0246528.
PR	08-NOV-2000; 2000US-0246532.
PR	08-NOV-2000; 2000US-0246509.
PR	08-NOV-2000; 2000US-0246610.
PR	08-NOV-2000; 2000US-0246611.
PR	08-NOV-2000; 2000US-0246613.
PR	17-NOV-2000; 2000US-0249207.
PR	17-NOV-2000; 2000US-0249208.
PR	17-NOV-2000; 2000US-0249209.
PR	17-NOV-2000; 2000US-0249210.
PR	17-NOV-2000; 2000US-0249211.
PR	17-NOV-2000; 2000US-0249212.
PR	17-NOV-2000; 2000US-0249213.
PR	17-NOV-2000; 2000US-0249214.
PR	17-NOV-2000; 2000US-0249215.
PR	17-NOV-2000; 2000US-0249216.
PR	17-NOV-2000; 2000US-0249217.
PR	17-NOV-2000; 2000US-0249218.
PR	17-NOV-2000; 2000US-0249244.
PR	17-NOV-2000; 2000US-0249245.
PR	17-NOV-2000; 2000US-0249264.
PR	17-NOV-2000; 2000US-0249265.
PR	17-NOV-2000; 2000US-0249297.
PR	17-NOV-2000; 2000US-0249299.
PR	17-NOV-2000; 2000US-0249300.
PR	01-DEC-2000; 2000US-0250160.
PR	01-DEC-2000; 2000US-0250391.
PR	05-DEC-2000; 2000US-0251030.
PR	05-DEC-2000; 2000US-0251988.
PR	05-DEC-2000; 2000US-0256719.
PR	06-DEC-2000; 2000US-0251477.
PR	08-DEC-2000; 2000US-0251856.
PR	08-DEC-2000; 2000US-0251866.
PR	08-DEC-2000; 2000US-0251869.
PR	08-DEC-2000; 2000US-0251989.
PR	11-DEC-2000; 2000US-0251990.
PR	05-JAN-2001; 2001US-0259678.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	
XX	Rosen CA, Barash SC, Ruben SM;
XX	WPI, 2001-488783/53.
DR	
XX	
PT	New nucleic acid molecules encoding 461 human secreted proteins for
PT	diagnosing, preventing, treating or ameliorating medical conditions and
PT	used as food additives or preservatives -
XX	
PS	Disclosure; SEQ ID No 1609; 980bp; English.
XX	
CC	The invention relates to isolated nucleic acid molecules and their
CC	encoded secreted proteins. The nucleic acids and proteins are used to
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC	are also used in diagnosing a pathological condition or susceptibility
CC	to a pathological condition. Antibodies to the proteins can also

CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence is a genomic DNA encoding a partial novel secreted protein of
CC the invention.

Query Match

1.0%; Score 27; DB 22; Length 2717;

Best Local Similarity 100.0%; Pred. No. 0.0076;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2149 tcacagaccacagagagacacacag 2175

Db 962 tcacagaccacagagagacacacag 988

Search completed: May 17, 2002, 09:32:07
Job time: 67704 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 13:38:43 ; Search time 3793.77 Seconds
(without alignments)
9861.842 Million cell updates/sec

Title: US-09-898-556A-3

Perfect score: 2772

Sequence: 1 cagggcgttaagctggtg.....ttctaccatctcaccct 2772

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 13736207 segs, 674847542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*

1: em_estda:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	721	26.0	1082	10	BM450709 AGENCOURT
2	677	24.4	1035	10	BM470740 AGENCOURT
3	585	21.1	1071	10	BE741389 601594255
4	581	21.0	590	9	AM245709 2823003.5
5	545	19.7	1009	10	BM449478 AGENCOURT
6	510	18.4	563	9	AI669533 w88h09.x
7	509	18.4	535	9	AL047439 DXFZP5861
8	492	17.7	538	9	AI802142 CX30C05.x
9	482	17.4	537	9	BE220142 hv67f07.x
10	471	17.0	624	9	BE178716 PM4-HT060
11	458	16.5	512	10	BE111754 7135a10.x
12	454	16.4	505	10	BE463501 hv24d02.x
13	449	16.2	486	9	AM401692 UT-HF-BK0
14	419	15.1	486	9	AV720986 AV720986
15	414	14.9	613	9	AM964787 EST376980
16	413	14.9	697	10	BE391140 601286705
17	405	14.6	673	10	BG928720 HMC72-1-B

c 18	402	14.5	453	9	AM513018	AM513018 xE76f03.x
c 19	399	14.4	481	9	AI276016	AI276016 GW08H02.x
c 20	395	14.2	451	9	AA450184	AA450184 Z442609.r
c 21	383	13.8	435	9	AI081665	AI081665 0063606.s
c 22	383	13.8	449	9	AI750087	AI750087 at35f07.x
c 23	378	13.6	455	10	BE501382	BE501382 hw31a02.x
c 24	378	13.6	728	10	BE746403	BE746403 601579525
c 25	372	13.4	424	9	AI016684	AI016684 0096605.x
c 26	372	13.4	428	9	AI440109	AI440109 t156d06.x
c 27	366	13.2	458	9	AI953623	AI953623 w423603.x
c 28	366	13.2	501	9	AM140124	AM140124 UT-H-B11-
c 29	363	13.1	426	9	AI828662	AI828662 t06f04.x
c 30	355	12.8	486	9	AI650983	AI650983 wa96d07.x
c 31	351	12.7	473	10	BF476173	BF476173 naa29c11.
c 32	350	12.6	385	9	AA218658	AA218658 z996d04.r
c 33	350	12.6	500	9	AI968444	AI968444 w265d09.x
c 34	345	12.4	551	10	BE395838	BE395838 601310077
c 35	344	12.4	508	9	AI860707	AI860707 w115d12.x
c 36	343	12.4	399	9	AI289399	AI289399 gw32605.x
c 37	340	12.3	392	9	AA218659	AA218659 z996d04.s
c 38	340	12.3	395	9	AI361705	AI361705 g218f04.x
c 39	337	12.2	513	9	AM958936	AM958936 EST371006
c 40	333	12.0	794	10	BI256434	BI256434 602974454
c 41	332	12.0	383	9	AA450118	AA450118 z442609.s
c 42	318	11.5	378	9	AA484745	AA484745 n881n06.s
c 43	315	11.4	369	9	AI159962	AI159962 q880f10.x
c 44	315	11.4	818	9	AI566746	AI566746 AL566746
c 45	312	11.3	501	10	R50617	R50617 yj57b01.r1

ALIGNMENTS

RESULT 1
BM450709 1082 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6394734 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494415
DEFINITION 5', mRNA sequence.
ACCESSION BM450709
VERSION 1 GI:18499749
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1082)
NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNAK12119 row: k column: 24
High quality sequence stop: 585.
location/Qualifiers
1. 1082
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5494415"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPO6; Site:1: NotI; Site:2: SalI. Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 Kb. Library constructed by Life Technologies."
BASE COUNT 298 a 261 c 292 g 226 t 5 others

ORIGIN

Query Match 26.0%; Score 721; DB 10; Length 1082;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 721; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 849 atcaaatatgaagagtttggccaggcttatacaaggatcaaaacctttagctccag 908
 DB 2 ATCAAAATATGAAGAGTTTGGCCAGGCTTATCAAGAGACTCAAACTCCTTAGCTCCAG 61
 QY 909 aagacaacaactgggagagacaccttatactatgtagtgggagagacttggcag 968
 DB 62 AAGACACAAACTGGGAGACACCTTACATGATACAGTGGGAGACAGCTTTGGCAGT 121
 QY 969 atgtcagtcctatcaaaaaccacaaagacacactctgggggaaagccttatgtgcag 1028
 DB 122 ATGTGAGTCTCATCAAAAACCAAGACACACTCTGGGGAAAGCTTATGTGTGACAG 181
 QY 1029 gaattgtggcagagctttagctgtgaagttcaaaacttgatcacatcagaagacactca 1088
 DB 182 GAATGTGGCGGAGGCTTACGTGAAGTCAAACTGATCACACATCAGAGACACACTCA 241
 QY 1089 ggggagaaacctatgtgtgtgaagattgtgacagagagcttacttgaagtcgaacctc 1148
 DB 242 GGGGAGAAACCTTATGTGTGACAGATGTGACAGAGGCTTACTTGGAGTGAACCTC 301
 QY 1149 ttacacatcagcgcagacacactcagagctcaagccttatgtgtgaaggaattgtggcag 1208
 DB 302 TTTCACATACACGCGACACACTCAGGGGTCAGGCTTATGTGTGACAGGAATGTGGCAG 361
 QY 1209 agctttagcctgaagtaaaacctcttaccacacagagggcgacacacttggggagaagcct 1268
 DB 362 AGCTTTAGCTTGAATCAAACTCATTAACCCACAGAGGGCCACACATCGGGAGAAAGCTT 421
 QY 1269 tatgttcagaggaattgtggcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1328
 DB 422 TATGTTTGCAGGAGATGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 481
 QY 1329 aggcacacatcagagagagagccttaccatltgtcagagaggtgtgacaaagcctttagcag 1388
 DB 482 AGGACACATTCAGGAGAGAGGCTTACATTTGACAGGAGTGTGACAGGCTTTAGCCAG 541
 QY 1389 aagtcacacctatcagacacactttagagacacacagagagagagagcctttagatgacaca 1448
 DB 542 AAGTACACCTCATCAAGACACTTAAGACACACAGAGAGAAAGCTTATGTATGTACACA 601
 QY 1449 gaattgtggcgttcaactttagctgtgaatcaaaacctcaaaacacacagagagacactca 1508
 DB 602 GAATGTGGCGCTCACTTTAGCTGGAATCAAACTCAAAACACACACAGAGACACACTCA 661
 QY 1509 ggggtttaaactttagctgtcctgtgagttgtgggagagtgctttagcctgaaagtcacacct 1568
 DB 662 GGGGTAAACCTTATGT 721
 QY 1569 a 1569
 DB 722 A 722

RESULT 2
 BM470740 1035 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT_6475358 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5579278
 DEFINITION 5', mRNA sequence.
 ACCESSION BM470740
 VERSION BM470740.1 GI:18519782
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1035)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: LNLN1235 row: k column: 23
 High quality sequence stop: 653.

FEATURES
 source
 1..1035
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5579278"
 /clone_id="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NCI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

BASE COUNT 259 a 249 c 268 g 255 t 4 others
 ORIGIN

Query Match 24.4%; Score 677; DB 10; Length 1035;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1655 caagcaccagagggagacacactcagggagagagcattttagtgcgtgagttgagcagag 1714
 DB 25 CACGACACAGAGAGACACCTCAAGGGAGAGACCATTTATGTGTGTGTGTGTGTGTGTGTGT 84
 QY 1715 cttaaatgataagtcacacctcaatttcacaccagagagacacattcagggagaaagcctt 1774
 DB 85 CTTTAAATGATTAAGTCCACCTCATTTTCACACAGAGACACATTTCAGGGAGAAAGCTTTT 144
 QY 1775 tatgtgcagagaggtgtgtgcagaaagtttcggcagaagcctttagttagcacaaag 1834
 DB 145 TATGTGACAGGAGTGTGACAGAAAGTTTGACAGAAAGCTTAAAGCTTAAAGCAAGAG 204
 QY 1835 ggcacacacaggtgccttgcgtgtgcagagaggtgtggaagcctttagttagtaagtaac 1894
 DB 205 GGCACACTCAGAGT 264
 QY 1895 tctcaatlaaacacacagagagacacagcaggggagagcctcaattgtgtcagagaggtgtg 1954
 DB 265 TCTCATTTAAACACAGAGAGACACAGCGAGGGGGAAGCTCATGTGTGTGTGTGTGTGTGT 324
 QY 1955 gcaaggtccttagccggcaggtacacacctcattagaacacagagagagacattagaagagaa 2014
 DB 325 GCAGAGCTTTAGCCGGAGTGCACACTCCTATTAGACACAGAGAGACATTAGAGAGAGAA 384
 QY 2015 gacctatattgcagaaaggtgtgacggggcctttagtcgagagtcacaccttaccagaca 2074
 DB 385 GCCTTATATTTGCAGAAAGGTGTGAGCGGGCTTATGTGTGTGTGTGTGTGTGTGTGTGTGT 444
 DB 445 TCAGAGGACACACACAGATAGAAACTTATGTATAGGGAAGTGTACAGCCTTTAG 504
 QY 2135 ccagagatctactctcagacacacagagagacacacagagtgctgtgtgtgtgtgtgtgtgtgt 2194
 DB 505 CCAGAGATCATTTCACTCAACACACAGAGACACACACAGTGTGTGTGTGTGTGTGTGTGTGT 564
 QY 2195 attctagatcacaaagtgtgagacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2254
 DB 565 ATTGCTGATACCAAAAGTGAGACATTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 624

QY 2255 agactgtatccatccacacaggaagtgtgctcatttccaggaagccctgccc 2314
|||||
Db 625 AGATTGTATTCATTCACATCGACATGAGAGAAATTCCTGGCTCATTTTCAGAGACCTGCCC 684
QY 2315 ttccctcactgtgagtgg 2331
|||||
Db 685 TTCTCCTACTGTGATGG 701
RESULT 3
BE741389 1071 bp mRNA linear EST 15-SEP-2000
LOCUS 601594255F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3948468 5',
DEFINITION mRNA sequence.
ACCESSION BE741389 GI:10155381
VERSION BE741389.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1071)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: L1CMB11 row: m column: 13
High quality sequence stop: 783.
Location/Qualifiers
1. 1071
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3948468"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 287 a 233 c 293 g 257 t 1 others
ORIGIN
Query Match 21.18; Score 585; DB 10; Length 1071;
Best Local Similarity 99.88; Pred. No. 7.5e-297;
Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1678 gggaggaagccattgtatgtcgtgagtgtgacagagcttaataagataagccacctca 1737
|||||
Db 36 GGGAGAAAGCCATTGTATGTCTGAGTGTGACAGAGCTTTATATCATTAAGTCCACCTCA 95
QY 1738 tttaacacagagacacatcagggggaagcctttatgtcagggagtgtgacagaa 1797
|||||
Db 96 TTTCACACACAGAGACACATTCAGGGGAAAGCCTTTATGTGACGAGAGTGTGACAGAA 155
QY 1798 gtttcgcaagaagcctaactgttttagcacaagaggcagacacacagtccttctgt 1857
|||||
Db 156 GGTTCGGCAGAGCCTTAACCTGTTTAGCACAAAGGGGACACCTCAGGTGCTTTGTGT 215
QY 1858 gcagggagtgtggtgcaaggtcttctgtaagtaactcattcaaacacagagagacac 1917

Db 216 GCAGGAGTGTGGGCAAGGCTTTTGTCTAGTTAACTTCATTAAACACAGAGACAC 275
|||||
QY 1918 acgagagggggaagccctcattgtgtcagggaggtgtggcaaggttaagccggagtcac 1977
|||||
Db 276 ACGCAGGGGGGAAAGCCCTCATGTGTGACAGGAGTGTGGCAAGGCTTTAGCCGGAGTCAC 335
QY 1978 acccattagacacacagagacacattcaagagagaagcccttatttgcagaagtgtg 2037
|||||
Db 336 ACCTCATTAGACACACAGAGACACATTCAGAGAGAAAGCCCTTATTTTGCAGAAAGTGTG 395
QY 2038 gacggggtcttaagtcggaagtcacaccccttaccagacacacagagacacactcagataga 2097
|||||
Db 396 GACGGGCTTTAGTCGGAAGTCACAACTTATCAGACATCAGACAGACACACTCAGATAGA 455
QY 2098 aacttatgtatagaggaaatgtgtacagcctttagccaggaatcactactatcagac 2157
|||||
Db 456 AACTTTATGTATAGGAAAGTGTACAGCCTTTAGCCAGAGCTCATCTTCATCAGAC 515
QY 2158 accagagacacacacagtcgtgtgcttttccagccattgctatagataccaaagtgaaga 2217
|||||
Db 516 ACCAGAGACACACACAGTGTGTGCTTTTCAGCCATTGCTAGTACCAAGTGAGAG 575
QY 2218 catctgtgtgtatgatgatgacagctgacgtgtgtgaagactgtatccatccacactg 2277
|||||
Db 576 CATCTGTGTGTGATTTATGATGACAGACTGTACTGTAGTACTGTATCTCATCTCACACTG 635
QY 2278 aagagaaatgtgtgtcatttccagagagccctgccc 2313
|||||
Db 636 AAGGAGAAATTCCTGCTCATTTTCAGGAGACCTGCC 671
RESULT 4
AN245709 590 bp mRNA linear EST 07-JAN-2000
LOCUS 2823003.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2823003 5',
DEFINITION mRNA sequence.
ACCESSION AN245709
VERSION AN245709.1 GI:6588702
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 590)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other ESTs: 2823003.3prime
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LNL) DNA Sequencing by: Berkeley MGC sequencing
project clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/db/rrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
trimming: cross_match from University of Washington Genome Center
PHRAP suite. Poly-T identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu
Plate: L1CMB10 row: O column: 4
High quality sequence stop: 582.
Location/Qualifiers
1. 590
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2823003"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 192 a 127 c 161 g 110 t

Query Match 21.0%; Score 581; DB 9; Length 590;
Best Local Similarity 100.0%; Pred. No. 8,7e-295;
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 544 gggcgggaatctctccacctggggaacactatccagaagaatcgaagaacagcagg 603
|||||
DB 10 GGGCAGGAATCCTCTCCACCTGGGMAACACTATCCAGATCAGAAACAGCAGG 69
OY 604 atccattctgctttagtggcaaaagcagaatgattcaagaaggaggaagctccagactcc 663
|||||
DB 70 ATCCATTCTGCTTTAGTGCAAGAGAGATTCAGAGGAGAGAGACTCCAGACTCC 129
OY 664 tttttgggagagtaagcaaaatggaactccaaggcacttccagcccaactgaagaac 723
|||||
DB 130 TTTTGGGAGACTTAAGCAAAATGGCACTTCAAGGCACCTTCCAGCCCACTTGAAAGAC 189
OY 724 aacagccacacagtcacaggaagaacacagagtgatgaatgaagtcagccctgaac 783
|||||
DB 190 AACAGCCACAGTCTCCAGAGACACACAGTGTGATTAAGGTCCAGCCCTGAC 249
OY 784 ggaaggcagatctagaagaacagaaatgattgattgattgaagtcagagattg 843
|||||
DB 250 GGAGGAGAGATCTAGAGAAACAGCAAAATGATGATGATGATGATGATGATGATG 309
OY 844 ggaagaatacaatataagaaggtttggcagagcttatacaaggatccaactccctagcc 903
|||||
DB 310 GGAATATCAAAATATGAGAGTTGGGCCAGCTTATCAAGAGATCAAACTCTTATAGCC 369
OY 904 tccagaagaacacaaactgggagagacactatgatactgaagtgggagagacacttg 963
|||||
DB 370 TCCAGAGACACAACTGGGAGACACTTACATGATACCTAGTGGGAGACAGCTTGG 429
OY 964 gcaagtatgcaatctcctcaacaaacccaaggacacactctgagggaaagccttattg 1023
|||||
DB 430 GCAGATGTCAGTCCATCAACAAACCCAGACACACTCTGGGGAAGCCTTATGTGT 489
OY 1024 gcaaggaaatggtggcgaagcttactgtaagttcaaacctgatacaatcagaagac 1083
|||||
DB 490 GCAGGAATGTGGGAGGCTTACGTGGAAGTCAAACTGATCAACATCAGAGAGAC 549
OY 1084 actcagggggaagaaccttattgtgcaagagattgagcga 1124
|||||
DB 550 ACTCAGGGAGAAACCTTATGTGTGCAAGATTTGTGACGA 590

RESULT 5
BM449478 1009 bp mRNA linear EST 05-FEB-2002
LOCUS BM449478
DEFINITION AGENCOURT_6401009 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493685
5', mRNA sequence.

ACCESSION BM449478
VERSION BM449478.1 GI:18498518
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNM at: <http://image.llnl.gov>
Plate: LLM12117 row: m column: 14
High quality sequence stop: 649.

FEATURES
source Location/Qualifiers

1..1009
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5493685"
/clone_1ib="NIH.MGC.67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

BASE COUNT 289 a 235 c 259 g 224 t 2 others

Query Match 19.7%; Score 545; DB 10; Length 1009;
Best Local Similarity 99.8%; Pred. No. 9.4e-276;
Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 206 ggcgtctggtgcatcagaagatggtgctggttacttaccacagagagtgagtgct 265
DB 196 GCGCTTCGTGGCATTCAGGAGATGCTGTGACTTCAACCCAGAGAGTGAGGCTTGT 255
OY 266 gagccctgtcagaagaccctgcagaaggagtgatgctggaacttaaccatctgt 325
DB 256 GAGCCCTGCTCAGAGACCTGACAGAGAGGTGATGCTGAGACTTAAACCATCTGT 315.
OY 326 ctcaatggaatctcatcttctcaacaaactctgctcagctggaagcgaaggagac 385
DB 316 CTCACATGGAATTCATCTCTTAAACAAACTCATTTGCTCAGCTGAGACGAGGAGAC 375
OY 366 gcccttgagagagagaaatgtccactggaactctgtccagaatcgaagcagaat 445
DB 376 GCCCTGAGAGAGAGAGAAATGTCCACTGGACCTCTGTCAGATGGAAGCCAGAAAT 435
OY 446 tcaacttagtcctctcctcctcctgattttctcagaagaagctccaagcaacatgt 505
DB 436 TCAACTTAGTCCCTCTGCTGCTGATTTTTCAGTCAAGCTCTCAGCAACATGT 495
OY 506 gtgagtgatcactctcctcagctgtttcaagttatcggcaggaatcctctccact 565
DB 496 GTGGGTAGTATCTCTCTCAGCTGTTTCAAGTTATGGGAGGAATTCCTCCACT 555
OY 566 gggaacacactatccagaagatcagaacaacagcagaatccatctgctttagtgcaa 625
DB 556 GGGAAACACATTCACAGAGATCAGAAACACAGCAGATCATCTCTTATGCGCAA 615
OY 626 agcagaatggaatcagaaggaggaagactccagactcctgtttgggaagtaagaaaaa 685
DB 616 AGCAGATGATTCAGAGGGGAGAACTCCAGACTCTCTTGGAGAGTAAGCAAAA 675
OY 686 tggcacttcaaaagcacttccagccaccttgaagaacaacagcagcagtcagaaga 745
DB 676 TGGCATTCAAAAGGACATCTTCAGCCCACTGAABAACAACAGCAGCAGCTCAAGA 735
OY 746 agacacacagtggtgatalataggttccagcccttgaagcgaagtgagagcttagagg 801
DB 736 AGACNACACAGTGTGATATAGGTCAGCCCTGAGGAGGAGGAGATCTAGAGG 791

RESULT 6

AI669533/c 563 bp mRNA linear EST 14-MAY-1999
LOCUS WB8H09.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2312801 3',
DEFINITION mRNA sequence.
ACCESSION AI669533 GI:4834307
VERSION AI669533
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 563)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/bdrrp/image/image.html
Seq primer: -400p from Gibco
High quality sequence stop: 454.
Location/Qualifiers
1. 563
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2312801"
/clone_1lb="NCI_CGAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneds
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Patricia Bonaldo."
BASE COUNT 162 a 125 c 150 g 126 t
ORIGIN
Query Match 18.4%; Score 510; DB 9; Length 563;
Best Local Similarity 99.8%; Pred. No. 2.5e-257;
Matches 560; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Y 2212 tggagacatctgctgctatgatcatgagacgtactgttaagactgtatccatc 2271
|||
Db 563 TGGAGACATCTGCTGATGATTGATGACAGACTGTAGTACTGATCTCCATC 504
Y 2272 caactgaagagaatctgctcatcttcagaagccctccctcctcactgtgagtg 2331
|||
Db 503 CACCTGAAGGAATGCTGCTCATTTTCAGGAGCCCTGCCCTTCCCTCAGTGTGATGG 444
Y 2332 tgggtctggaaccgcgctcaggtatagatgagcaagagcagtcacatgcccagcca 2391
|||
Db 443 TGGGTGTGGAAACCCGGTCAAGTAATGATGAGGAGGAGCAATGCCCCAGGCA 384
Y 2392 gatagggttgggtactcctgctgaaccccaacctlaagctgaagaacgtccggcctaaatc 2451
|||
Db 383 GATAGGGGTGGGTACTGCTGTAACCCCAACTTAAAGCTGAAGACAGTCCGGGGCAATTC 324
Y 2452 ctcaatagaattgagaacctgctctccatcttgctgtgcttccctccgattgatccaa 2511
|||
Db 323 CTCATACTGAATGAGAACTGTCTCCATTGGTGTGCTTCTCCGATTGATGCCAA 264

Y 2512 ccctcaactattttacgtatcacctgcctctcctaattgttttaactgctgccc 2571
|||
Db 263 CCCTTCACACTATTTTACGTATACCTGCCCTTTCTTAATGTGTTTACACGCTGTCGCC 204
Y 2572 acccttgatgctgctcttgcatcactacaatcagtcacgtgtatccctcattctg 2631
|||
Db 203 ACCTTTGAGTGGTGGCTTTGTCATATCTTACAAATCAGTCAACGTATTCCTCCATTCCTG 144
Y 2632 agcccaataaagaccagctcagctcagctgagtgagagaataaccctgctgagaggt 2631
|||
Db 143 AGCCCATAAAGACCCAGACACTCAGAGAGAGAAATACCCCTGCTGGGGGT 84
Y 2692 tgggagacactccctgcgtacccctcctcactcagagagctgttcttgctaataaatc 2751
|||
Db 83 TGGGACCACTCCCTGATACCCCTCTCAGAGAGCTGTCTTCTCATTAATTC 24
Y 2752 ttctaccatccctcaccct 2772
|||
Db 23 TTTCTACCCATCTCTCACCCCT 3
RESULT 7
AL047439 535 bp mRNA linear EST 29-FEB-2000
LOCUS AL047439
DEFINITION DKFZ58610320_r1 586 (synonym: hute1) Homo sapiens cDNA clone
ACCESSION AL047439
VERSION AL047439.1 GI:4727354
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 535)
Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Wambutt, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Wambutt R
MPS
Am Kioferpitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No SI sequence available.
This clone (DKFZ58610320) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. 535
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZ58610320"
/clone_1lb="586 (synonym: hute1)"
/tissue="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
BASE COUNT 159 a 128 c 143 g 105 t
ORIGIN
Query Match 18.4%; Score 509; DB 9; Length 535;
Best Local Similarity 100.0%; Pred. No. 8.2e-257;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 1071 catagaagacactcagggagagaacctatctgtgcaagagttgtgagcagagcttt 1130
|||
Db 1 CATCAGAGACACACTCAGGGGAGAAACCTTATGTGTGTCAGAGATTGTGGACGAGGCTTT 60
Y 1131 acttgaagtcgaacctttacacatcagcgagacacactcagaggtcgaagccttatgtg 1190

Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL, send email to: info@image.lnl.gov

Seq primer: -400P from GIBCO
High quality sequence stop: 450.

FEATURES

source

Location/Qualifiers

1..537

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3178501"

/clone_id="NCI-CGAP-Lu24"

/tissue_type="carcinoid"

/lab_host="DH10B"

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Plasmid DNA from the normalized library NCI-CGAP-Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonalido."

BASE COUNT 150 a 118 c 144 g 125 t
ORIGIN

Query Match 17.4%: Score 482; DB 9; Length 537;
Best Local Similarity 99.8%: Pred. No. 1.4e-242;

Matches 532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2236 gcatgagactgtactgtgaagactgtatctccaccctgaagagaattgtctgtc 2295
|||||
DB 537 GCATGAGACTGTACTGTGAAGACTGTATCTCCATCCACCTGAAGAGAAATGTGGCTC 478

QY 2236 attttcagaagccttgccttccctcactgtgagtggtgtgttgaaaccggtcagt 2355
|||||
DB 477 ATTTTCAGAGACCCCTGCTTCCTCCACGTGATGCTGGTGTGAAACCCGCTCAGT 418

QY 2356 aatgtagtgcagagagcactcaaatgcccagagagatagggtgttactgtgtaa 2415
|||||
DB 417 AATGATAGTGGCAGAGGAGGAGTCCAGGAGATAGGGGTGGTACTGGTGAAA 358

QY 2416 cccaacttaagctgaagacagctcccgctaaatccctcatataatgaagacctgtc 2475
|||||
DB 357 CCCAACCTTAAAGCTGAAGACAGTCCGGCTAAATCTCATATCTGAATGAACACTGTC 298

QY 2476 ttcccatlgtgtgtcttctcctcgatgatcccaacccttaacctatttaataacc 2535
|||||
DB 297 TTCCCATTTGGTGTCTTCTCCCATGATGCCAACCTTCACCTATTTTACGTATACC 238

QY 2536 tgccttccctcaattgttttaacactgtgtgcccaccttttggtgtgtgtgtat 2595
|||||
DB 237 TGCCCTTTCCTAATGTGTTTACACTGTGTGCCACCTTTTGAGTGTGCTTTGCAAT 178

QY 2596 acttaacaatcagtcacagtgatccctcatctcgaagcccaataaagaccagactcag 2655
|||||
DB 177 ACTTAACAATCAGTCAAGCTGTATTCCTCATTTTCGAGGCCCTAATAAGCCAGACTCAG 118

QY 2656 ctgcaagtgaagagaatcacacctgtctgtgaggttgaggagacacctcctgtcactccct 2715
|||||
DB 117 CTGCAGTGAGGAGAAATCACCCCTGCTGTGGGGTGTGGGGACCACTCCCTCATCCTCCT 58

QY 2716 ctccactgagagctgtcttctgtcctaataatctttttaccatcctccta 2768
|||||
DB 57 CTCACATGAGAGCTGTCTTTGCTCAATTAATTTCTTTTACCACTCCTCA 5

RESULT 10

BE178716

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

MEDLINE

COMMENT

JOURNAL

QY 974 agtccatcaaaaacccaagacacactctggggaagccttatgtgtgcaaggatg 1033
|||||
Db 249 ACTCCATCAAAAACCCCAAGGACACACTGGGGGAAAGCCTTATGTGCGAGGAATG 308
QY 1034 tggggaagccttacgtggaagtaacactgtatcaacatcagagacacactcaaggga 1093
Db 309 TGGGCGAGCCTTACGTGAAGTCAAACTGATCACACACGAGAGACACACTCAGGGGA 368
QY 1094 gaaacctatgtctgcaagatgtgtgagagagccttacttgaagtcaacctcttac 1153
Db 369 GAAACCTTATGTGTGCAAGATTTGTGACGAGCCTTTACTTGAAGTCCAACTCCTTAC 428
QY 1154 acatcaagcgagacacactcagggtcaggccttatgtgtgcaagaaatgtctggaagcctt 1213
Db 429 ACATCAGCGAGACACACTCAGGCGCTCAAGCCTTATGTGCAAGAAATGTGGGAGAGCTT 488
QY 1214 tagctgaagtcataacactatccacacagagggcgacactggggaagccttatgt 1273
Db 489 TAGCCTGAAGTCAAACTCATTACCCACAGAGGGCGCACTGGGAGAGCCTTATGT 548
QY 1274 ttgcaaggaaatgtggcgctgtcttgcagaga 1306
Db 549 TTGCAAGGAATGTGGCGCTTGTGCGCAGCA 581

RESULT 11
BF111754 512 bp mRNA linear EST 20-OCT-2000
LOCUS BF111754.1
DEFINITION 7155a10.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:3523123 3', mRNA sequence.
ACCESSION BF111754
VERSION BF111754.1 GI:10941444
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 512)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 447.
Location/Qualifiers
1. 512
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3523123"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site: 1: Not 1; Site: 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and as circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NbHSF pool 1:
309384-147935, 323208-325895 Soares Nb2HP pool 1:
145032-147935, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NbHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NbHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 144 a 111 c 140 g 117 t

ORIGIN

Query Match 16.5%; Score 458; DB 10; Length 512;
Best Local Similarity 99.8%; Pred No. 6.6e-230;
Matches 508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2260 tgaatccatccacactgaagagaatgtgtgtcattttcaaggagcccttccct 2319
|||||
Db 512 TGTATCTCCATCCACTGAGAGGAAATTTGCTGGCTATTTTCAGAGAGCCCTCCCT 453
QY 2320 caactgtgagatgtgtgtgtgtgtgaaacccggtcaggtatgatatgtgcagaggaagta 2379
|||||
Db 452 CACTGTGATGTGTGGGTTGTGAAACCCGCTCAGGTAAATGATGTGGCAGAGGCGAGTCA 393
QY 2380 aatgtcccaagcagataagggtgtgtgtacactgtgtgaaacccaacttaagcttaagacagt 2439
|||||
Db 392 AATGCCAGGAGATPAGGGGTGGGTACTGTGTGAACCCCAACCTTAAACTGAAGACAGT 333
QY 2440 ccgagctaaatccataactgaatgaagacctgtcttccattgtgtgtcttcc 2499
|||||
Db 332 CCCGGCTAAATCTCATACTGAATTGAGAACCTGTCTTCCATTTGTGTGCTTCCCTCC 273
QY 2500 gatgatcccaacacctcaacctattttagtataacctgtcccttccctaatgtgttttac 2559
|||||
Db 272 GATGATCCCAACCCCTTCACCTATTATTAAGTATACCTGCGCTTCTTAATTTGTTTTAC 213
QY 2560 actgctgtgcccacctttttagtgtgtccttgcatacttaataatgaagcgtgat 2619
|||||
Db 212 ACTGCTGTGCCACCTTTTGAAGTGTGCTTGTGCATTAATCAATCAAGTCAAGCTGTAT 153
QY 2620 tcccatatctagaccatataaagaccagactcagctcagtgagagagaataacacc 2679
|||||
Db 152 TCCCTATTTCTGAGCCCAATAAAGACCAAGACTGAGCTGAGAGAGAAATCACCC 93
QY 2680 tgcgttgaaggttggggaccactccctgcattcccttccactgagagctgtctttgc 2739
|||||
Db 92 TGCCTGTGGGGTGGGGACCACTCCCTGCATCCCTTCACAGAGAGCTGTTCTTTGCG 33
QY 2740 tcaataaatctttcttaccatccctca 2768
|||||
Db 32 TCAATAAATTTCTTTTTCATCCATCTCA 4

RESULT 12
BE463501/c 505 bp mRNA linear EST 27-JUL-2000
LOCUS BE463501/c
DEFINITION h24d02.x1 NCI-CCAP_Ki611 Homo sapiens cDNA clone IMAGE:3183843 3',
mRNA sequence.
ACCESSION BE463501
VERSION BE463501.1 GI:9509274
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 505)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL, send email to:
info@image.lnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 455.

FEATURES
source
Location/Qualifiers
1.505
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:318343"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and M.1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 142 a 110 c 140 g 113 t
ORIGIN

Query Match 16.4%; Score 454; DB 10; Length 505;
Best Local Similarity 99.8%; Pred. No. 8.4e-228;
Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2263 atctccatccactggaaggaattgctgctcatcttcaggaagcctgcctccctccac 2322
|||
DB 505 ATCTCCATCCACTGGAAGGAATGCTGCTCATTTTCAGAGACCTGCCCTTCTCCAC 446
|||
QY 2323 tctgagatgctggttctggaaccggtcagtaatatagtgaggaagcagtcacaa 2382
|||
DB 445 TGTGAGATGCTGGGTGTGGAACCCGGTCAAGTAATGATGTGCGAGAGCACTCAAT 386
|||
QY 2383 gcccaagcagatagggtgtgtactctgtgaaaccaacttaagctgaagacagtc 2442
|||
DB 385 GCCCAGGCAATAGGGGTGGGTACTGTGTAACCACTTAAGCTGAAGACAGTCC 326
|||
QY 2443 ggtctaaatccatcagtaattgaaacctgtctccattgtgtgcttccctccat 2502
|||
DB 325 GGCTAAATCTCATACGAATTGAAACCTGTCTTCCCATTTGGTGTGCTTCTCCGAT 266
|||
QY 2503 tgatcccaaccttcacctaattacgtatccctccctcccttaattggttttaact 2562
|||
DB 265 TGAATCCCAACCTTCACCTAATTTACGTATACCTGCCCTTCCATATGGTTTTACACT 206
|||
QY 2563 gctgtgccacacctttgagtggtgcttgcataacttaacaaatcagtcacgtatcc 2622
|||
DB 205 GCTGTGCCACCTTTGAGTGTGCTTGTGCACTTACAATCAATCAAGTCACTATATCC 146
|||
QY 2623 cctatcttgagccataaagaaccagactcagtcagtgaggaagagaataaccctgc 2682
|||
DB 145 CCTATTGTGAGCCATTAAGAACCCAGACTCAGTGCAGTGAGAGAAATCAACCTGC 86
|||
QY 2683 tctggaagcttgaggagaccacccctgcataccctcctccacagagagctgttctgtca 2742
|||
DB 85 TGTGGGGGTGTGGGGACACACTCCCTGCATCCCTCTCCACAGAGAGCTGTTTGTGCTCA 26
|||
QY 2743 ataaatcttcttaccacccctc 2767
|||
DB 25 ATAAATCTTTCTACCCATCCCTC 1
|||

RESULT 13
AM401692 486 bp mRNA linear EST 16-FEB-2000
LOCUS UI-BRO-aaf-c-03-0-UI.r1.NIH.MGC_36 Homo sapiens cDNA clone
DEFINITION IMAGE:3053549 5', mRNA sequence.
ACCESSION AM401692
VERSION AM401692.1 GI:6920378
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS 1 (bases 1 to 486)
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/dbp/image/image.html
Seq primer: M13 forward.

FEATURES
source
Location/Qualifiers
1.486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3053549"
/clone_lib="NIH-MGC_36"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTI)"
/note="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (0.5-1.5kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 138 a 113 c 136 g 98 t 1 others
ORIGIN

Query Match 16.2%; Score 449; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 3.6e-225;
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 963 ggcagatgtcagtcctcatcaaaaaccagaagacactctgggggaagccttatgtg 1022
|||
DB 8 GGCAGTATGTCAGTCCATCAAAACCAAGGACACACTGTGGGGAAACCTTATGTG 67
|||
QY 1023 tgcagggaatgtgggagagcttcaagtgagtcgaacctgatacacatcagaagaca 1082
|||
DB 68 TGCAGGAATGTGGCGGCTTTACGTGGAAGTCAAACTGATCACATCAGAGACA 127
|||
QY 1083 cactcaggggaacacttatgtgtcgaagatgtgtgacgagagccttacttgaagtcg 1142
|||
DB 128 CACTCAGGGGGAACCTTATGTGCAAGATTTGTGACGAGCTTACTTGGAAATCG 187
|||
QY 1143 aaactcttacacatcagggagacactcaggtcgaaccttatgtgtcgaagaaatgt 1202
|||
DB 188 AACCTTTTACACATCAGCGGACACACTCAGGGCTCAAGCTTATGTGTGCAAGGAATGT 247
|||
QY 1203 gggcagagccttaagcctgaagtcgaacctccttccacagaggggagcagagggag 1262
|||
DB 248 GGGCAGAGCTTTACCTGAAGTCAAACTCAATTCACCAAGAGGGCCACACTGGGAG 307
|||
QY 1263 aagccttatgttgcaggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1322
|||
DB 308 AAGCCTTATGTTTGCAGGGAATGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 367
|||
QY 1323 cacagaagacacatcagaagaagccttacatttcagaagaagtgtgagcaagcctt 1382
|||
DB 368 CACAAGAGGACACATTCAGAGAGAGGCTTACATTTTCAGAGAGTGTGAGCAAGGCTTT 427
|||
QY 1383 agccagaagtacacacctatcagacactt 1411
|||
DB 428 AGCCAGAAGTCAACACTCATCAGACACTT 456
|||

RESULT 14

AV720986
 LOCUS AV720986 486 bp mRNA linear EST 16-OCT-2000
 DEFINITION AV720986 HTB Homo sapiens cDNA clone HTBIC05 5', mRNA sequence.
 ACCESSION AV720986
 VERSION AV720986.1 GI:10818138
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 486)
 AUTHORS Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.
 TITLE Homo sapiens cDNA HTB clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
 Tel: 86-21-50801919 (ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 FEATURES
 source
 1. 486
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HTBIC05"
 /clone_lib="HTB"
 /issue_type="Hypothalamus"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 104 a 134 c 110 g 137 t 1 others
 ORIGIN
 Query Match 15.1%; Score 419; DB 9; Length 486;
 Best Local Similarity 100.0%; Pred. No. 2.4e-209;
 Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2269 atccaactgaaggagaatgctgctcaatlttcaggagccctgcctcctcactgctga 2328
 Db 1 ATCCACCTGAGAGGAATTCGCTGCTCATTTTCAGAGAGCCCTGCCCTTCCTACATGTGA 60
 QY 2329 tgggtgggtgtgggaaccggctgagatagtagtgagcaggagcagtcgaatgtccag 2388
 Db 61 TGGTGGGTTGGGAACCCGGTCAAGTATAGTAGTGGCAGAGGCACTCAAAATGCCAG 120
 QY 2389 gcaagataagggtgggtacctgtgtgaaccacaacttaagctgaagacagtcgccgctaa 2448
 Db 121 GCAGTAGGGGTGGGTACTGTGTGAACCAACCTTAAGCTGAAGACAGTCCCGCTAA 180
 QY 2449 atcctcatctgaattggaacactgctctcccaattggtgtgtcttccctcagattgaccc 2508
 Db 181 ATCTCATACTGAATGGAACCTGTCTCCCATTTGGTGTCTTCTCCGATGTGATCC 240
 QY 2509 caaccctcacctatttaagtatactacgccccttcccaattgtgttttaacgtgtgtg 2568
 Db 241 CAACCTTACCTATTTATAGTATACCTGCCCTTTCCTAAATGGTTTATACCTGCTGTG 300
 QY 2569 cccacctttagtgtgtccttgcatactacaatcagtcacaagtcgtatctccctatt 2628
 Db 301 CCCACCTTTTGTAGTGTGCTTGTGATCTTACAATAGTCAACGATATATCCCTATT 360
 QY 2629 ctgagccataaagaaccagactcagctcagtgagtgaggaatcaaccctgtctgtg 2687
 Db 361 CTGAGCCCATTAAGACCCAGACTCAGCTCAGTGAGGAGAGAAATCACCTGCTGTGG 419

RESULT 15
 LOCUS AM964787/c 613 bp mRNA linear EST 01-JUN-2000
 DEFINITION AM964787 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM964787
 VERSION AM964787.1 GI:8154743
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 613)
 AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharep, S., Gaspard, R., Gay, C., Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J., and Quackenbush, J.
 TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
 JOURNAL Unpublished (2000)
 COMMENT Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 207
 Seq primer: Forward.
 FEATURES
 source
 1. 613
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGH"
 /note="Vector: pBluescriptskm"
 BASE COUNT 171 a 140 c 158 g 144 t
 ORIGIN
 Query Match 14.9%; Score 414; DB 9; Length 613;
 Best Local Similarity 99.8%; Pred. NO. 1.1e-206;
 Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2223 tgtgtgtattatgcatagagactgactgctgaagctgtatccatccacactgaaga 2282
 Db 544 TGTGTGTATATATGATAGATAGACTGTACTGTGTAAGCTTGTATCTCATCCACCTAAGA 485.
 QY 2283 gaattgtcgtcatttccaggagccctgcctcctcactgactgtgagtggtgtgtga 2342
 Db 484 GAATGTGCTGCTCATTTTCAGAGAGCCCTGCCCTTCCTCAGTGTGATGTGGTGTGA 425
 QY 2343 aaccgggtcagtgatagtagtgagcaggagcagtcgaatgccagcagataggggtg 2402
 Db 424 AACCGGTCAAGTATATGATAGTGGCAGAGGACAGTCAATATGCCAGCATAGAGGGTGG 365
 QY 2403 gtacctgtgaaccacaacttaagctgaagacagtcgccgctaaatccatatactgaa 2462
 Db 364 GTACCTGTGAACCAACCTTAAGCTGAAGACAGTCCCGCTTAATCTCATACTGANA 305
 QY 2463 ttggaacactgctctcccatctgtgtgtcttccctcagattgacccaactcaacta 2522
 Db 304 TTGAACACCTGTCTCCCATTTGGTGTCTTTCCTCCGATGTATGCCAACCTTCACTA 245
 QY 2523 tttaacgataccgtgccccttcccaattgtgttttaacgtgtgtgtgtgtgtgt 2582
 Db 244 TTTTATAGTATACCTGCCCTTTCCTAAATGGTTTATACCTGCTGCCCCACCTTTTGTG 185
 QY 2583 ggtgccttgcatactacaatcagtcacaagtcgtatccctattctgagccataaa 2642
 Db 184 GTTGCTTTGATATCTTACAATCAGTCAACGTGTATTCCTTATTTTGGCCCATATAA 125
 QY 2643 gaccagactcagctcagtgagtgaggaatcaaccctgtctgtg 2687
 Db 124 GACCCAGACTCAGCTGAGTGAAGAGAAATCACCTGCTGTGG 80

Sat May 18 14:46:07 2002

us-09-898-556a-3.rst

Page 11

Search completed: May 17, 2002, 09:22:49
Job time: 71046 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic : nucleic search, using sw model

Run on: May 17, 2002, 13:55:18 ; Search time 105.07 seconds
(without alignments)
6480.404 Million cell updates/sec

Title: US-09-898-556A-3

Perfect score: 2772

Sequence: 1 cagcgcgttaagctgttg.....ttctaccatcctaccct 2772

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
4: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
5: /cgn2_6/prodata/1/lna/PCrUS_COMB.seq:*
6: /cgn2_6/prodata/1/lna/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	0.8	2382	4	US-09-588-256-9
2	20	0.7	1892	2	US-08-933-750C-66
3	20	0.7	1892	3	US-09-234-613-66
4	19	0.7	3240	4	US-09-262-773-7
5	19	0.7	3240	4	US-09-262-773-3
6	19	0.7	3264	4	US-09-262-773-5
7	19	0.7	3268	4	US-09-262-773-1
8	19	0.7	3810	2	US-08-475-844-8
9	19	0.7	3810	5	PCr-US95-08429-8
10	19	0.7	20137	4	US-09-262-773-206
11	19	0.7	20138	4	US-09-262-773-9
12	19	0.7	23071	4	US-09-262-773-210
13	19	0.7	246240	2	US-08-724-394A-20
14	19	0.7	246240	2	US-08-724-394A-21
15	19	0.7	246240	2	US-08-724-394A-22
16	18	0.6	575	1	US-08-507-016-8
17	18	0.6	1027	4	US-09-303-524A-1
18	18	0.6	1558	1	US-08-416-870C-9
19	18	0.6	2133	2	US-08-820-170A-11
20	18	0.6	2133	3	US-09-055-699-11
21	18	0.6	2133	4	US-09-273-565-11
22	18	0.6	2133	4	US-09-565-538-11
23	18	0.6	3754	2	US-08-820-170A-12
24	18	0.6	3754	3	US-09-055-699-12
25	18	0.6	3754	4	US-09-273-565-12
26	18	0.6	3754	4	US-09-565-538-12
27	18	0.6	7886	2	US-08-751-189-2

c	28	18	0.6	7886	2	US-09-060-836-2	Sequence 2, Appl
c	29	18	0.6	7886	4	US-09-184-445-2	Sequence 2, Appl
c	30	17	0.6	80	5	PCT-US91-02942-99	Sequence 9, Appl
c	31	17	0.6	273	3	US-08-956-182-21	Sequence 21, Appl
c	32	17	0.6	289	1	US-08-592-126-124	Sequence 124, App
c	33	17	0.6	347	4	US-08-905-223-133	Sequence 133, App
c	34	17	0.6	532	4	US-09-328-111-648	Sequence 648, App
c	35	17	0.6	563	4	US-09-385-982-433	Sequence 433, App
c	36	17	0.6	732	3	US-08-956-182-16	Sequence 16, Appl
c	37	17	0.6	1362	4	US-09-171-461-47	Sequence 47, Appl
c	38	17	0.6	1561	1	US-07-662-007B-38	Sequence 38, Appl
c	39	17	0.6	1561	1	US-07-968-971A-1	Sequence 1, Appl
c	40	17	0.6	1561	1	US-07-824-247-38	Sequence 38, Appl
c	41	17	0.6	1561	1	US-07-824-247-41	Sequence 41, Appl
c	42	17	0.6	1561	1	US-08-142-473A-1	Sequence 1, Appl
c	43	17	0.6	1561	1	US-08-424-406-1	Sequence 1, Appl
c	44	17	0.6	1561	1	US-08-464-523B-5	Sequence 5, Appl
c	45	17	0.6	1561	1	US-08-469-203A-1	Sequence 1, Appl

ALIGNMENTS

```
RESULT 1
US-09-588-256-9
; Sequence 9, Application US/09588256
; Patent No. 6291665
; GENERAL INFORMATION:
; APPLICANT: Gaffney, Thomas
; APPLICANT: Flavler, Albert
; APPLICANT: Gates, Krista
; APPLICANT: Wendland, Juergen
; APPLICANT: Ayad-Durieux, Yasmina
; APPLICANT: Dietrich, Fred
; APPLICANT: Philippsen, Peter
; TITLE OR INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-30908A
; CURRENT APPLICATION NUMBER: US/09/588,256
; CURRENT FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: Asbyda gossypii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2382)
; US-09-588-256-9

Query Match 0.8%; Score 21; DB 4; Length 2382;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1990 ggccttgctgaactaact 1895
1875 ggccttgctgaactaact 1895
|||||
Db 1990 ggccttgctgaactaact 2010

RESULT 2
US-08-933-750C-66
; Sequence 66, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
```

```

1 TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
2 NUMBER OF SEQUENCES: 98
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: Incyte Pharmaceuticals, Inc.
5 STREET: 3174 Porter Drive
6 CITY: Palo Alto
7 STATE: CA
8 COUNTRY: USA
9 ZIP: 94304
10
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Diskette
13 COMPUTER: IBM Compatible
14 OPERATING SYSTEM: DOS
15 SOFTWARE: FASTSEQ for Windows Version 2.0
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/933 750C
18 FILING DATE: September 23, 1997
19 CLASSIFICATION: 336
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER:
22 FILING DATE:
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Billings, Lucy J.
25 REGISTRATION NUMBER: 36,749
26 REFERENCE/DOCKET NUMBER: PF-0356 US
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 415-855-0555
29 TELEFAX: 415-845-4166
30 TELEX:
31
32 INFORMATION FOR SEQ ID NO: 66:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 1892 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: single
37 TOPOLOGY: linear
38 IMMEDIATE SOURCE:
39 LIBRARY: BRSTMOT03
40 CLONE: 641127
41
42 US-08-933-750C-66
43
44 Query Match 0.7%; Score 20; DB 2; Length 1892;
45 Best Local Similarity 100.0%; Pred. No. 3,4;
46 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
47
48 QY 1419 cacacagagagaagcctta 1438
49 ||||||||||||||||||
50 DB 1289 CACACAGAGAGAACCTTA 1308
51
52 RESULT 3
53 US-09-234-613-66
54 Sequence 66, Application US/09234613
55 Patent No. 6132973
56 GENERAL INFORMATION:
57 APPLICANT: Lal, Preeti
58 APPLICANT: Hillman, Jennifer L.
59 APPLICANT: Bandman, Olga
60 APPLICANT: Shah, Purvi
61 APPLICANT: Au-Young, Janice
62 APPLICANT: Yue, Henry
63 APPLICANT: Guegler, Karl J.
64 APPLICANT: Corley, Neil C.
65 TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
66 NUMBER OF SEQUENCES: 98
67 CORRESPONDENCE ADDRESS:
68 ADDRESSEE: Incyte Pharmaceuticals, Inc.
69 STREET: 3174 Porter Drive
70 CITY: Palo Alto
71 STATE: CA
72 COUNTRY: USA
73 ZIP: 94304
74
75 COMPUTER READABLE FORM:

```

```

1 MEDIUM TYPE: Diskette
2 COMPUTER: IBM compatible
3 OPERATING SYSTEM: DOS
4 SOFTWARE: FASTSEQ for Windows Version 2.0
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: US/09/234,613
7 FILING DATE:
8 CLASSIFICATION:
9 PRIORITY APPLICATION DATA:
10 APPLICATION NUMBER: US/08/933,750
11 FILING DATE: September 23, 1997
12 ATTORNEY/AGENT INFORMATION:
13 NAME: Billings, Lucy J.
14 REISTRATION NUMBER: 36,749
15 REFERENCE/DOCKET NUMBER: PF-0356 US
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: 415-855-0555
18 TELEFAX: 415-845-4166
19 TELEX:
20 INFORMATION FOR SEQ. ID NO.: 66:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 1892 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: single
25 TOPOLOGY: linear
26 IMMEDIATE SOURCE:
27 LIBRARY: BRSTNOT03
28 CLONE: 641127
29
30 US-09-234-613-66

```

Query Match	0.7%	Score 20	DB 3	Length 1892
Best Local Similarity	100.0%	Pred. No. 3.4		
Matches	20	Conservative	0	Mismatches
			0	Indels
			0	Gaps
Qy	1419	cacacagagagaagcctta	1438	
Db	1289	cacacagagagaagcctta	1308	

```

: RESULT 4
: US-09-262-773-7
: Sequence 7, Application US/09262773
: Patent No. 6225451
: GENERAL INFORMATION:
: APPLICANT: Ballinger, Dennis G.
: APPLICANT: Ding, Wei
: APPLICANT: Wagner, Susanne
: APPLICANT: Hess, Mark A.
: TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
: FILE REFERENCE: Myriad 3
: CURRENT APPLICATION NUMBER: US/09/262,773
: CURRENT FILING DATE: 1999-03-04
: NUMBER OF SEQ ID NOS: 210
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 7
: LENGTH: 3240
: TYPE: DNA
: ORGANISM: human
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (98)..(523)
: US-09-262-773-7

```

Query Match	0.7%	Score 19	DB 4	Length 3240
Best Local Similarity	100.0%	Pred. No. 11		
Matches 19	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 2068	tcagacatcagagagacaca	2086		
Db 1559	tcagacatcagagagacaca	1577		

RESULT 5
US-09-262-773-3
; Sequence 3, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3244
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(2017)
US-09-262-773-3

Query Match 0.7%; Score 19; DB 4; Length 3244;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2068 tcagacatcagagagacaca 2086
|||||

DB 1563 tcagacatcagagagacaca 1581

RESULT 6
US-09-262-773-5
; Sequence 5, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3264
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(496)
US-09-262-773-5

Query Match 0.7%; Score 19; DB 4; Length 3264;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2068 tcagacatcagagagacaca 2086
|||||

DB 1583 tcagacatcagagagacaca 1601

RESULT 7

US-09-262-773-1
; Sequence 1, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3268
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(2041)
US-09-262-773-1

Query Match 0.7%; Score 19; DB 4; Length 3268;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2068 tcagacatcagagagacaca 2086
|||||

DB 1587 tcagacatcagagagacaca 1605

RESULT 8
US-08-475-844-8
; Sequence 8, Application US/08475844
; Patent No. 5972643
; GENERAL INFORMATION:
; APPLICANT: Lobanenko, Victor V.
; APPLICANT: Neiman, Paul E.
; APPLICANT: Klenova, Elena M.
; APPLICANT: Goodwin, Graham H.
; APPLICANT: Filippova, Galina N.
; APPLICANT: Collins, Steven J.
; TITLE OF INVENTION: CTCF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend Kourlie and Crew
STREET: One Market Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,844
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,680
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14538A-11-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600

```
TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3810 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   ORIGINAL SOURCE:
;     ORGANISM: Homo sapiens
;     IMMEDIATE SOURCE:
;     CLONE: human CTCF cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 292..2475
;   FEATURE:
;     NAME/KEY: exon
;     LOCATION: 281..1074
;   OTHER INFORMATION: /label= exon2
;   FEATURE:
;     NAME/KEY: exon
;     LOCATION: 1075..1245
;   OTHER INFORMATION: /label= exon3
;   FEATURE:
;     NAME/KEY: exon
;     LOCATION: 1246..1379
;   OTHER INFORMATION: /label= exon4
;   FEATURE:
;     NAME/KEY: exon
;     LOCATION: 1380..1499
;   OTHER INFORMATION: /label= exon5
;   FEATURE:
;     NAME/KEY: exon
;     LOCATION: 1500..1649
;   OTHER INFORMATION: /label= exon6
;   FEATURE:
;     NAME/KEY: exon
;     LOCATION: 1650..1810
;   OTHER INFORMATION: /label= exon7
;   FEATURE:
;     NAME/KEY: exon
;     LOCATION: 1810..1992
;   OTHER INFORMATION: /label= exon8
;   OTHER INFORMATION: /label= exon8
US-08-475-844-8

Query Match      0.7%; Score 19; DB 2; Length 3810;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1755 cattcagggaagaacctt 1773
DB 1492 cattcagggaagaacctt 1510

RESULT 9
PCT-US95-08429-8
; Sequence 8, Application PC/TUS9508429
; GENERAL INFORMATION:
;   APPLICANT:
;   TITLE OF INVENTION: CTCF
;   NUMBER OF SEQUENCES: 21
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patent Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US95/08429
;     FILING DATE: 15-JUN-1995
;     CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 08/261,680
```

```
FILING DATE: 17-JUN-1994
; CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Parmelee, Steven W.
;     REGISTRATION NUMBER: 31,990
;     REFERENCE/DOCKET NUMBER: 14538A-11-1PC
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 206-467-9600
;     TELEFAX: 415-543-5043
;   INFORMATION FOR SEQ ID NO: 8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 3810 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     ORIGINAL SOURCE:
;       ORGANISM: Homo sapiens
;       IMMEDIATE SOURCE:
;       CLONE: human CTCF cDNA
;     FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 292..2475
;     FEATURE:
;       NAME/KEY: exon
;       LOCATION: 281..1074
;     OTHER INFORMATION: /label= exon2
;     FEATURE:
;       NAME/KEY: exon
;       LOCATION: 1075..1245
;     OTHER INFORMATION: /label= exon3
;     FEATURE:
;       NAME/KEY: exon
;       LOCATION: 1246..1379
;     OTHER INFORMATION: /label= exon4
;     FEATURE:
;       NAME/KEY: exon
;       LOCATION: 1380..1499
;     OTHER INFORMATION: /label= exon5
;     FEATURE:
;       NAME/KEY: exon
;       LOCATION: 1500..1649
;     OTHER INFORMATION: /label= exon6
;     FEATURE:
;       NAME/KEY: exon
;       LOCATION: 1650..1810
;     OTHER INFORMATION: /label= exon7
;     FEATURE:
;       NAME/KEY: exon
;       LOCATION: 1810..1992
;     OTHER INFORMATION: /label= exon8
;     OTHER INFORMATION: /label= exon8
PCT-US95-08429-8

Query Match      0.7%; Score 19; DB 5; Length 3810;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1755 cattcagggaagaacctt 1773
DB 1492 cattcagggaagaacctt 1510

RESULT 10
US-09-262-773-206
; Sequence 206, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;   APPLICANT: Ballinger, Dennis G.
;   APPLICANT: Ding, Wei
;   APPLICANT: Wagner, Susanne
;   APPLICANT: Hess, Mark A.
;   TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
```

;; TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;; FILE REFERENCE: Myriad 3
;; CURRENT APPLICATION NUMBER: US/09/262,773
;; CURRENT FILING DATE: 1999-03-04
;; NUMBER OF SEQ ID NOS: 210
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 206
;; LENGTH: 20137
;; TYPE: DNA
;; ORGANISM: human
US-09-262-773-206

Query Match 0.7%; Score 19; DB 4; Length 20137;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2068 tcagacatcagagagacaca 2086
|||||
Db 14680 tcagacatcagagagacaca 14698

RESULT 11
US-09-262-773-9
; Sequence 9, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 20138
; TYPE: DNA
; ORGANISM: human
US-09-262-773-9

Query Match 0.7%; Score 19; DB 4; Length 20138;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2068 tcagacatcagagagacaca 2086
|||||
Db 14681 tcagacatcagagagacaca 14699

RESULT 12
US-09-262-773-210
; Sequence 210, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 23071
; TYPE: DNA

;; ORGANISM: human
US-09-262-773-210

Query Match 0.7%; Score 19; DB 4; Length 23071;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2068 tcagacatcagagagacaca 2086
|||||
Db 17614 tcagacatcagagagacaca 17632

RESULT 13
US-08-724-394A-20/C
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereeto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Filts, Renee A. 35,136
; REGISTRATION NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note="HLA-H.CONTIG"
US-08-724-394A-20

Query Match 0.7%; Score 19; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2298 tttagagagcctgcctt 2316
|||||
Db 36115 tttagagagcctgcctt 36097

RESULT 14
US-08-724-394A-21/c
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match 0.7%; Score 19; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2238 ttccagagccctgccctt 2316
|||||
DB 36115 TTTCAGAGCCTGCCCTT 36097

RESULT 15
US-08-724-394A-22/c
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.

APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

Query Match 0.7%; Score 19; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2238 ttccagagccctgccctt 2316
|||||
DB 36115 TTTCAGAGCCTGCCCTT 36097

Search completed: May 17, 2002, 16:27:42
Job time: 9144 sec

Sat May 18 14:46:06 2002

us-09-898-556a-3.rni

Page 7

